

BUILDING DETECTION AND REGULARISATION USING DSM AND IMAGERY INFORMATION

YOUSIF A. MOUSA^{ab*} (y.mousa@postgrad.curtin.edu.au)

PETRA HELMHOLZ^b (petra.helmholz@curtin.edu.au)

DAVID BELTON^b (d.belton@curtin.edu.au)

DIMITRI BULATOV^{ac} (dimitri.bulatov@iosb.fraunhofer.de)

^a*Spatial Sciences, Curtin University, Perth, Australia*

^b*Civil Engineering Department, Al-Muthanna University, Al-Muthanna, Iraq*

^c*Fraunhofer IOSB, Ettlingen, Germany*

*Corresponding author

Abstract

An automatic method for the regularisation of building outlines is presented, utilising a combination of data- and model-driven approaches to provide a robust solution. The core part of the method includes a novel data-driven approach to generate approximate building polygons from a list of given boundary points. The algorithm iteratively calculates and stores likelihood values between an arbitrary starting boundary point and each of the following boundary points using a function derived from the geometrical properties of a building. As a preprocessing step, building segments have to be identified using a robust algorithm for the extraction of a digital elevation model. Evaluation results on a challenging dataset achieved an average correctness of 96.3% and 95.7% for building detection and regularisation, respectively.

KEYWORDS: building detection, DEM extraction, digital surface model, laser scanning, polygonal simplification, regularisation

INTRODUCTION AND PREVIOUS RESEARCH

THE MAPPING OF BUILDING OUTLINES is relevant for numerous applications such as urban planning, three-dimensional (3D) city modelling and map updating. As the manual digitisation of building outlines is expensive and time-consuming, the aim of several approaches is to automate the process starting with the raw sensor data. A typical pipeline for the regularisation of building outlines consists of three crucial steps: (i) computation of the digital elevation model (DEM – sometimes termed a digital terrain model or DTM) from the digital surface model (DSM); (ii) building mask detection; and (iii) building outlining. The following paragraphs refer to previous work achieved on these steps as well as their insufficiencies.

Extracting a DEM is the key step for building detection (Mongus et al., 2014) and it must be computed first (Bulatov et al., 2014). The existing DEM extraction algorithms are

either based on “raw” point clouds or on a rasterised DSM where each pixel has a specific value which corresponds to a height. Methods based on point clouds can be difficult and time-consuming when searching for adjacent points in border regions, while working on a raster DSM can overcome such problems (Meng et al., 2009b). In addition, 3D point density is not always sufficient, which can lead to a negative impact on the quality of the results (He et al., 2014; Rottensteiner et al., 2014). Interpolating point cloud data into a raster may help to overcome this issue because it offers enough points that are consistently sampled in a regular grid.

Based on the literature, DEM extraction algorithms can be described as slope-based, linear-prediction-based and methods based on morphological filtering (Sithole and Vosselman, 2004; Liu, 2008; Mongus et al., 2014). In addition, directional scanning filters have also proved to be promising algorithms for DEM generation (Meng et al., 2009b, 2010; Perko et al., 2015; Mousa et al., 2017).

Slope-based methods are based on the assumption that slope angles between ground points are distinctly lower than slopes between ground and non-ground points. Several methods have been developed (Axelsson, 2000; Vosselman, 2000; Sithole, 2001; Shan and Sampath, 2005). The selection of a suitable slope threshold is critical because the terrain slope can be non-uniform, even in the same scene, depending on terrain complexity. Adaptive slope thresholding is therefore implemented (Sithole, 2001) to handle such limitations. Promising results can be achieved with such methods in flat areas, while their efficiency decreases with increasing terrain slope (Liu, 2008; Mongus et al., 2014).

Linear prediction or interpolation-based methods start by estimating a rough terrain. Then, height differences or residuals between points and the estimated rough terrain are minimised by linear least squares interpolation. Negative residuals are given higher weights than positive ones assuming that the estimated roughness is usually interpolated over the actual ground surface. A number of approaches have been modified based on this concept (Kraus and Pfeifer, 1998; Lee and Younan, 2003). Cubic-spline surface minimisation can be also used because of its robustness against outliers (Bulatov and Lavery, 2010; Bulatov et al., 2014). However, extraction of detailed ground surfaces and small objects might be difficult by using such interpolation methods (Sithole and Vosselman, 2004; Liu, 2008).

Classical morphological filtering methods have been commonly implemented for DEM extraction (Kilian et al., 1996; Zhang et al., 2003). The idea is based on applying morphological operators, such as erosion and dilation (Haralick et al., 1987), in greyscale images using a structural element (SE). The size of the SE is critical in eliminating buildings with diverse sizes; Zhang et al. (2003) therefore proposed progressive filtering by gradually increasing the sizes of the SE. The local elevation difference and slope within the size of the SE were applied to identify ground points. Chen et al. (2007) applied a similar approach, but used an adaptive slope threshold. Nevertheless, percentile rank filters (for example, 5%, 10%, 20%, and 40%) were used to mitigate the effect of outliers. More recently, Mongus et al. (2014) also applied a morphological filtering approach for building detection purposes in three steps: (i) smoothing the DSM using a Gaussian kernel was applied as preprocessing step; (ii) objects having smooth or continuous surfaces were labelled as ground objects, while other objects were removed; and (iii) the input laser points were classified as ground points based on their elevation differences from DEM and the slope gradient.

Directional scanning methods have also attracted attention. Meng et al. (2009b) proposed a multi-directional ground filtering (MGF) approach based on lidar-DSM. The basic idea was to apply scanline filtering, from left to right and right to left, within a

moving window of pre-defined dimensions. Points were classified as ground or non-ground based on both the height difference with the lowest point in the conducted scanlines and on the slope with their adjacent points. Selection of the lowest point in sloping areas is the main limitation of this method. Perko et al. (2015) overcame this limitation by considering the local slope of the terrain; however, smoothing of the DSM is required which may lead to possible inaccuracies.

Points belonging to a DEM can be identified based on the following physical characteristics: (i) lowest elevation in a local area; (ii) slope angle; (iii) elevation difference; and (iv) surface homogeneity (Meng et al., 2010). The majority of existing algorithms have been built on these characteristics. The slope is the most sensitive parameter in the process of DEM generation (Zhang and Whitman, 2005). In addition, defining a slope threshold in terms of terrain information in the analysed scene is, to some extent, subjective (Zhang et al., 2003). This problem becomes obvious in high-resolution DSMs because the slope value between neighbouring pixels is considerable without transitioning from ground to non-ground, or vice versa (Mousa et al., 2017).

After the extraction of the DEM, the so-called normalised DSM (nDSM), representing buildings, trees and other non-ground features, can be generated by subtracting the DEM from the DSM. For the purpose of generating a building map, the problem is related primarily to removing trees. To do so, the normalised difference vegetation index (NDVI) can be utilised if it is available. Otherwise, the multi-return properties of laser points (Meng et al., 2009a) and/or analysing the height difference of a point and its neighbours, for example, using a planarity measure (West et al., 2004), are possible options. If a comprehensive classification is required, convolutional neural networks (CNNs) (Long et al., 2015) can be utilised; these aim to train a parametric system learning object identification jointly with a classifier (Marcos et al., 2018) in an end-to-end method. Its advantage is in its effectiveness in image classification with pixel-level accuracy. The main disadvantage, however, lies in an expensive training effort. Nevertheless, an nDSM is an essential input in processing (for example, as used by Marmanis et al. (2018) and Piramanayagam et al. (2018)) which can only be computed after extracting the DEM. The generated building map can be either building regions (such as a binary image) or segmented laser points. In both cases, a list of boundary points can be obtained using the Moore contour-tracing algorithm (Gonzalez et al., 2004) or convex-hull-based procedures (Pohl et al., 2017).

The approaches used for building outline regularisation can be grouped into three categories: (i) model-driven (Sohn et al., 2012; Brédif et al., 2013); (ii) data-driven (Pohl et al., 2017); or (iii) a combination of both models (He et al., 2014). The goal of these methods is to reduce the number of building boundary points to the minimum required to describe the shape of the building. More specifically, the result from model- and data-driven approaches is to find a polygon describing the same properties of the original building object with a significant reduction in the number of given boundary points.

Model-driven approaches rely on several preselected parametric building models to be fitted with a given boundary-point dataset; they are therefore more robust against noise. Such approaches are based on the orthogonality characteristics of building outlines (right-angled corners), which is true for the majority of existing buildings but can be a challenge for complex buildings. Therefore, non-rectangular buildings can be incorrectly simplified or be represented by overly complex shapes (Avbelj, 2015). The most common model-driven approaches for building outlining are implemented using a minimum bounding rectangle (MBR) procedure (Gerke et al., 2001; Dutter, 2007; Arefi, 2009; Kwak and Habib, 2014; Avbelj, 2015). Overcoming the problem of noisy data is the main advantage of model-

driven approaches; the main drawback is that only buildings with right angles can be modelled, which is not always the case for complex buildings.

In contrast to model-driven approaches, data-driven approaches do not require prior knowledge of the building shape; they therefore offer greater flexibility to model any building shape. Nevertheless, data-driven approaches still face difficulties in the case of missing data to be recovered (Sohn et al., 2012) and it is often difficult to formulate constraints imposing regularity (Kwak and Habib, 2014). The most common data-driven approaches can be classified, in terms of the underlying method, into three groups: (i) the Douglas–Peucker (DP) algorithm (Douglas and Peucker, 1973); (ii) the random sampling consensus (RANSAC) method (Fischler and Bolles, 1981); and (iii) the Hough transform (Hough, 1962). Many of the existing data-driven approaches have utilised an initial solution based on the DP algorithm (Maas and Vosselman, 1999; Wang et al., 2006; Jwa et al., 2008; Sohn et al., 2012; He et al., 2014) because it is easy to implement and is able to maintain the original shape (Song and Miao, 2016). However, the efficiency of the DP algorithm decreases with increasing irregularity of the building boundary points. This is because the algorithm is inherently sensitive to both noise and the position of the starting point. In addition, building characteristics such as angle detection and area preservation are not considered in the processing. Therefore, a more robust data-driven approach is preferable.

The final step for the building outline regularisation is an adjustment that aims to find the best-fitting polygon edges with respect to the input boundary points for non-rectilinear buildings. For rectilinear buildings, orthogonality constraints must be added. If the polygon edges are adjusted subsequently, the topological relationships between such edges will be lost (Avbelj, 2015). Many of the existing procedures tend to find the dominant building direction, which then is used as the basis to adjust the other polygon edges accordingly (Arefi, 2009; Awrangjeb, 2016; Höhle, 2017). However, the challenge is to adjust all parameters simultaneously (Avbelj, 2015). Therefore, the best fitting of polygon edges must be achieved jointly with orthogonality constraints to provide the best solution.

Considering the drawbacks of existing methods, as summarised above, this study proposes a workflow with the aims of:

- (1) Employing a robust DEM extraction algorithm that does not need a slope parameter nor the smoothing of the DSM.
- (2) Developing a new data-driven approach to generate approximate polygons for both rectilinear and non-rectilinear buildings.
- (3) Combining the robustness of model-driven approaches with the flexibility of data-driven approaches in order to build a comprehensive strategy to deal with the building outline regularisation problem more effectively.
- (4) Solving regularisation of rectilinear buildings by implementation of orthogonality constraints and the best fitting of boundary points with respect to their corresponding edges simultaneously in the adjustment process.

The paper is structured as follows. In the next section the methodology is introduced. The results are then presented and evaluated, with the last section containing conclusions.

METHODOLOGY

DSM refinement is the first step in the proposed workflow presented in Fig. 1. Then, the DSM is separated into a DEM and an nDSM. Further, trees are removed using NDVI image and planarity data to generate a building mask. The next step is to combine

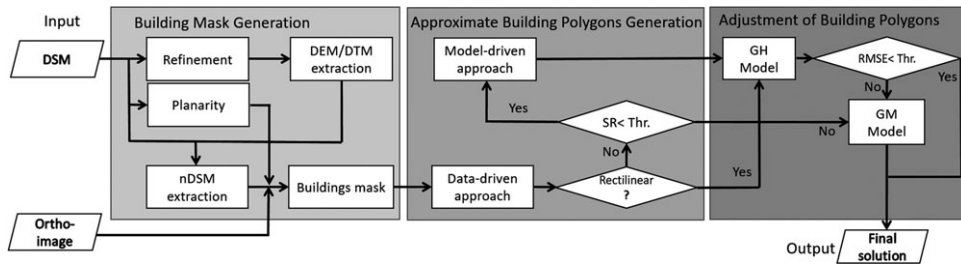


FIG. 1. Workflow of the proposed methodology.

the data-driven approach with the model-driven approach to obtain approximate building polygons. Buildings satisfying the rectilinearly conditions (for example, an even number of vertices) will be processed via a least squares adjustment using the Gauss–Helmert (GH) model. Otherwise, they are assumed to be non-rectilinear and will be processed through the Gauss–Markov (GM) model to provide the final solution of the building outline regularisation.

Building Mask Generation

As a DSM refinement step, a median filter (*MF*) is applied. While small elevated objects are eliminated in this step, it allows the modelling of the building outlines to be more precise, as well as improving the DSM quality against outliers (Bulatov et al., 2014). The size of the *MF* must be chosen carefully because closely adjacent buildings could be merged to a single complex shape. This could result in an increase in the percentage of false positives and, as a negative consequence, make the regularisation result become less reliable. The DEM is generated using the network of ground points (NGPs) algorithm (Mousa et al., 2017). Then, the nDSM is created by subtracting the DEM from the DSM in order to extract non-ground points. Next, the non-ground points other than buildings (primarily high vegetation) have to be removed from the nDSM to generate the building mask. To do so, an NDVI mask is created from the ortho-image. The NDVI values range from -1 to 1 , with pixels having NDVI values greater than 0.15 being the most likely to belong to vegetation. Non-vegetation objects usually having a low NDVI value (less than about 0.1). However, when vegetation is located in shadow areas, their NDVI values are very low and therefore may not be removable. To overcome this limitation, the authors considered the planarity measure (West et al., 2004). Considering a neighbourhood of a 3D point, a so-called *structure tensor* is obtained whose eigenvalues provide several important measures. The most popular is the planarity measure which, essentially, assesses how the neighbourhood of a point can be approximated by a plane. To compute the structure tensor, it is helpful to interpret the slightly smoothed DSM as a point cloud. Based on the cylinder-like extraction of neighbours proposed by Gross and Thoennessen (2006), the planarity measure is particularly suitable for differentiating between the ground and flat roofs (value close to 1), areas around walls (value of 0 , because there is a jump in elevation), trees (values close to 0 , except for the crown, which constitute smaller regions) and sloping roofs (which, depending on inclination angle, have an approximately constant value between 0 and 1). The overall decision rule using NDVI and planarity measure is given by:

$$\text{BuildingsMask} = (\text{nDSM} > \text{threshold and } (\text{NDVI} < 0.15 \text{ and planarity} > 0.6)). \quad (1)$$

Pixels satisfying equation (1) are labelled as building pixels. The result is a binary image separating all building segments from the background. In Fig. 2, the marked point (red cross) belongs to a tree in a shadow area and was selected as an example. Its NDVI value is 0.08 which is too low to be eliminated by the NDVI threshold, which is set to 0.15. However, its planarity measure is 0.26, and therefore lower than the required threshold of 0.6, indicating that this point does not belong to a building. Therefore, the point is correctly eliminated from the nDSM when forming the building mask. It has to be emphasised that the parameters in equation (1) were determined empirically, but the robustness of the approach – that nDSM, NDVI and planarity are sufficient to discard trees – is confirmed because few changes occur by modifying these thresholds by around 10%. In addition, small objects (for example, with an area of less than 10 m^2) are eliminated because: firstly, they are most likely to be associated with tree crowns in shadow, vehicles or just noise; and secondly, buildings (or parts of buildings) with an area less than 10 m^2 are so small they rarely exist. Finally, morphological filtering is performed to fill small holes that may occur.

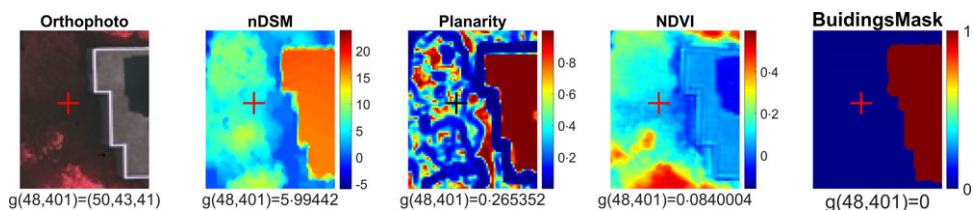


FIG. 2. Generation of the building mask. Left to right: orthophoto, nDSM, planarity map, NDVI mask and building mask. The red cross belongs to a tree in a shadow area.

Approximate Building Polygon Generation

Data-driven Approach to Generate an Approximate Building Polygon. The goal of this step is to find the minimum number of vertices, from a list of boundary points, that represents a building outline as similar as possible to its original building shape. These boundary points are collected in Cartesian coordinates for each building segment by applying the Moore contour-tracing algorithm (Gonzalez et al., 2004).

A list of consecutive boundary building points is given. Let $P = \{p_1, p_2, \dots, p_I\}$, where I is the number of boundary points. The goal is to find $V = \{v_1, v_2, \dots, v_J\}$ with $V \subset P$ and $3 \leq J \leq I$, where J is the number of vertices and V is a closed circular sequence of J vertices ($V_{J+1} = V_1$) (Backes and Bruno, 2013). In this context, the orthogonal distance error ($dist$) between each boundary point (i) and its corresponding line segment (j), as well as the root mean square error (RMSE), are calculated as shown in equation (2):

$$\text{RMSE}(V_j) = \left(\frac{\sum_i \text{dist}(p_i, \overrightarrow{VV_j})^2}{I^+} \right)^{\frac{1}{2}} \quad (2)$$

where $j = 2, \dots, J$, $\text{dist}(p_i, \overrightarrow{VV_j})$ is the orthogonal distance (see the red lines in Fig. 3) from the line connecting V and V_j to the boundary point p_i , V is the previously determined vertex V_{j-1} while I^+ denotes the number of points i running from V_{j-1} to V_j .

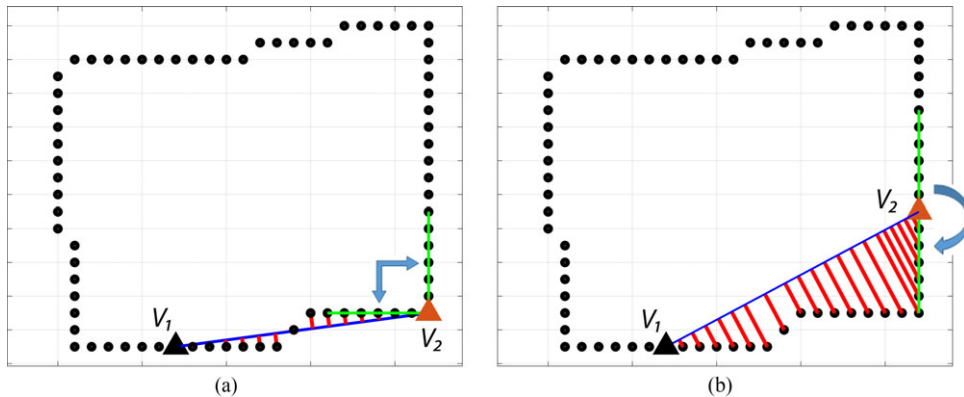


FIG. 3. Examples of conducted straight lines (blue lines) from the start point (black triangle) to the following ones. The red lines are the orthogonal distances. The green lines represent a distance of 1.5 m (6 pixels with a 0.25 m GSD) for the angle-detector threshold at the point under evaluation (orange triangles).

Fig. 3 shows an example of a list of boundary points with a 0.25 m ground sample distance (GSD), which needs to be simplified. Starting from an arbitrary point V_1 , for example, the one shown as a black triangle, the aim was to find the locations of the most probable vertices. Logically, points located at a building corner (the point presented by the orange triangle in Fig. 3(a)) can be nominated as a vertex (V_2). However, finding this point automatically is a challenging task. To do so, the following criteria should be considered:

- (1) Distance (D): V_2 can be defined as the farthest point from V_1 , and at the same time achieves three criteria: acceptable RMSE, preservation of the building's area and it (V_2) should be located on a corner in order to avoid unnecessary boundary points.
- (2) Error (E): the RMSE of the boundary points between V_1 and V_2 must be acceptable, and must therefore lie under a pre-defined threshold.
- (3) Area (A): the calculated area of the points which should be as close as possible to the original area while reducing the unnecessary points between V_1 and V_2 .
- (4) Angle (Θ): V_2 must not be located on a straight line considering the previous and subsequent points; there is an angle even for curves as the aim is to find the corner points of the building. The priority is given to angles close to 90° .

These four criteria seem to be the most essential with respect to the physical characteristics which should be considered to solve the problem of building simplification through a list of given points. However, formulating these characteristics in one mathematical model solution is a challenging task. Based on these characteristics a likelihood equation consisting of three major terms is formulated as follows:

$$l(i) = A + aD|\sin(\Theta)| - bE^2 \quad (3)$$

where l is the likelihood for a boundary point i to be nominated as a vertex, and a and b are weight factors to balance the equation.

The first term on the right-hand side of equation (3) represents the calculated area (A) after reducing the number of redundant boundary points. Sometimes, the calculated area is larger than the original area, which typically occurs when the interior angles of a building exceed 180° . In this case, the calculated area has to be modified.

The second term in equation(3) combines the distance (D) and angle (θ) hypotheses. In Fig.3 the blue lines represent the distance D from the starting point (black triangles) to the point under evaluation (orange triangles), while θ is the calculated angle at the point under evaluation. A high weight for this term is required ($a = 20$ was determined by empirical testing and fixed in all the tests) to make sure that the distance between two adjacent vertices is maximised within an acceptable RMSE. This is a reasonable assumption if the evaluated boundary point is close to a building corner. Otherwise, lower weight should be assigned to the distance, especially when the evaluated point is located on a straight line. For this reason, this weight a is multiplied by the absolute sine of the angle at the point under evaluation. The angle is calculated by using the boundary points before and after the current point. Fig.3 shows the distances for the angle detector (AD) depicted by the two green lines (6pixels in length) intersecting at the point under evaluation. When the point under evaluation is located at a building's corner (Fig.3(a)), the angle is close to 90° . In contrast, when the point under evaluation is located on a straight line (Fig.3(b)), the angle is close to 180° . Since $|\sin(90^\circ)| = |\sin(270^\circ)| = 1$ (a maximum) while $|\sin(180^\circ)| = 0$ (a minimum), the evaluated point attracts greater weight when it is located at a corner while it is assigned zero weight when it is located on a straight line.

The last term in equation(3) represents the mean square error (MSE) of the orthogonal distances, visualised by the red lines in Figs. 3(a) and (b) from the constructed straight (blue) lines. The orthogonal distance for each boundary point, along with the constructed line (E), is calculated according to equation (2). This is squared and multiplied by the weight ($b = 2$, which was determined by empirical testing and fixed in all tests).

The process recursively constructs straight lines from the starting point to each subsequent boundary point, as shown in Fig.3. On each occasion, the likelihood value is computed according to equation(3) and indexed. This process stops when reaching the maximum of the pre-defined RMSE threshold. Then, the boundary point which has the highest likelihood value will be nominated as a vertex. The process is repeated, starting from the last accepted vertex, until returning to the first nominated vertex. Boundary points that achieved the highest likelihood value for each individual process will be considered as the most likely vertices. In the results section it will be demonstrated that, contrary to the DP algorithm, the proposed method does not show an exaggerated dependence on the starting point.

Model-driven Approach for Buildings not Passing the Rectilinearity Conditions. In this work, only the first level of detail of the MBR is applied to detect whether a building has a rectangular shape. The decision as to whether a building has rectangular shape or not is made by thresholding the so-called *similarity ratio* (SR):

$$SR = (A_s/A_m) \quad (4)$$

where A_s is the area of the building segment (the yellow area surrounded by black boundary points in Fig.4(d)) and A_m is the area of the derived MBR as presented by the red box in the same figure. Kwak and Habib (2014) considered this ratio to evaluate their reconstruction result and called it the *area ratio*. This approach is applied to cope with buildings having high levels of noise at their border, caused by inadequate data or removing vegetation covering building roofs. Therefore, the data-driven approach may not

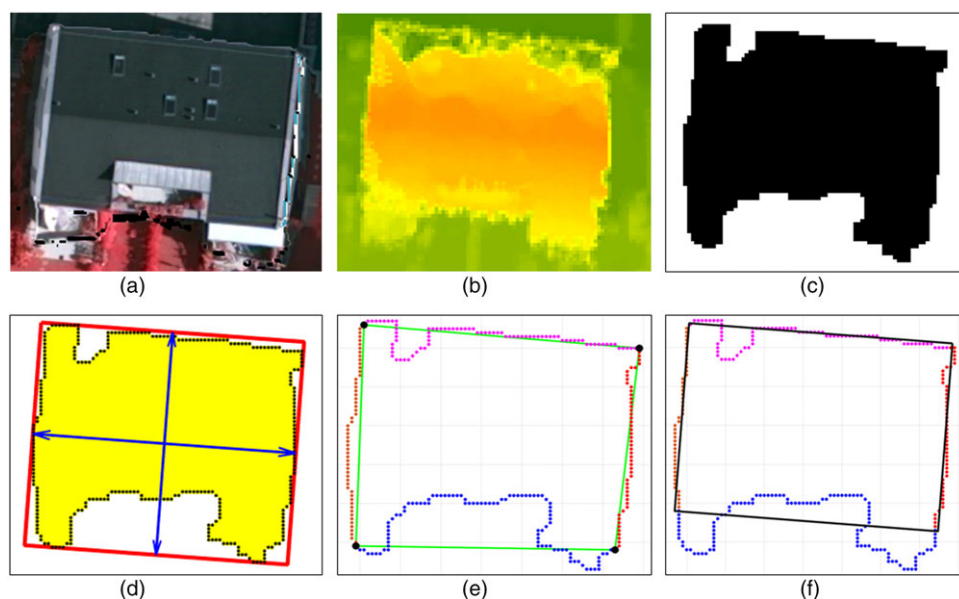


FIG. 4. Building simplification procedure based on the similarity ratio (SR). (a) Orthophoto showing a building with its reference. (b) DSM data. (c) Building segment. (d) MBR. (e) Initial outline. (f) Final outline after Gauss-Helmert processing.

always provide the appropriate initial polygons for such buildings. Fig. 4 demonstrates an example for a rectangular building, represented by its reference polygon in Fig. 4(a) and DSM data in Fig. 4(b); the building has a very noisy border as shown in Fig. 4(c).

The SR value is always close to one when a building is rectangular, while it becomes smaller with decreasing probability of being a rectangle. If this value is within the range of a preselected threshold, a decision of being a rectangle will be made. Then, an approximate polygon with only four vertices is introduced, as depicted by the black dots and connected by the green lines in Fig. 4(e); in this figure all boundary points have been assigned different colours according to their corresponding edges. The vertices have been selected from the boundary points as they have the minimum distance to the corners of the bounding box. Lastly, processing is done by GH: the final building outlines are presented as black solid lines associated with up-to-date labelling of the boundary points in Fig. 4(f). While the advantage of the SR value is that it can overcome a high level of noise in the boundary points more efficiently, the disadvantage is that small details in building outlines can disappear.

Adjustment of Building Polygons

After obtaining approximate polygons for the building outlines, model adjustment is the final step in the workflow. When the approximated polygons fulfil the rectilinear conditions, the GH model is applied. Otherwise, buildings are assumed to be non-rectilinear and will be processed with the least squares fitting approach based on the GM model (Avbelj, 2015) by modifying the given weights to the observation equations.

Gauss–Helmert (GH) Model. Orthogonality means that each pair of connected line segments must have a right angle at their connection point. In a mathematical sense, if a line segment j has a normal vector $\mathbf{n}_j = (n_x, n_y)$, then the next line segment must have a perpendicular normal vector $\mathbf{n}_{j+1} = (n_y, -n_x)$. All parallel line segments have the same normal vector, as highlighted by the red arrows in Fig. 5.

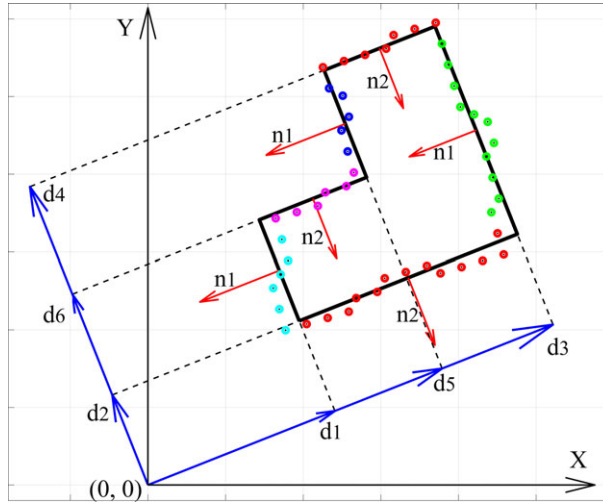


FIG. 5. Orthogonality functional model. Boundary points (dots) coloured differently according to their corresponding edges. Red arrows indicate the direction of normal vectors. The blue arrows are the distances d_i computed from the origin. The black solid line is the final rectilinear polygon after the adjustment.

The observation equations for a boundary point i , expressed by its Cartesian coordinates (x_i, y_i) located on its corresponding line segment, are given as follows:

$$f = n_x x_i + n_y y_i - d_j = 0 \quad (5)$$

$$f = n_y x_i - n_x y_i - d_{j+1} = 0 \quad (6)$$

where d_j and d_{j+1} are the distances from the origin of the coordinate system parallel to, and perpendicular to, the extension of a line segment j (blue arrows in Fig. 5). The unknown parameters (u) are n_x, n_y and d_1, \dots, d_J where J is the number of vertices or line segments. Therefore, the functional model can be written as follows:

$$F(\hat{\mathbf{u}}, \hat{\mathbf{l}}) = 0 \quad (7)$$

where $\hat{\mathbf{u}}$ and $\hat{\mathbf{l}}$ are the unknowns and the observations, respectively. The normals n_x and n_y must have a length equal to one and satisfy the following constraint:

$$g = n_x^2 + n_y^2 - 1 = 0. \quad (8)$$

Because the unknowns and observations for a point in the line segment observation equation cannot be separated, and each constraint contains more than one observation, the mixed

model or GH model adjustment has to be applied (Skaloud and Lichti, 2006). While the GH adjustment was implemented by Avbelj (2015), in this work it is modified as outlined below.

The model is linearised for the approximated values \mathbf{u}^o of the unknowns \mathbf{u} and the observations \mathbf{l} as follows:

$$F(\hat{\mathbf{u}}, \hat{\mathbf{l}}) = F(\mathbf{u}^o, \mathbf{l}) + \left. \frac{\partial F}{\partial \mathbf{u}} \right|_{\mathbf{u}^o, \mathbf{l}} (\hat{\mathbf{u}} - \mathbf{u}^o) + \left. \frac{\partial F}{\partial \mathbf{l}} \right|_{\mathbf{u}^o, \mathbf{l}} (\hat{\mathbf{l}} - \mathbf{l}) + \varepsilon = 0 \quad (9)$$

$$\mathbf{A}_{I,u} \delta_{u,1} + \mathbf{B}_{I,2l} \mathbf{v}_{2l,1} + \mathbf{w}_{I,1} = 0 \quad (10)$$

where $\mathbf{A} = \frac{\partial \mathbf{F}}{\partial \mathbf{u}}$ is the design matrix of partial derivatives of the observation equations with respect to the unknowns at the approximated values \mathbf{u}^o ; $\delta = (\hat{\mathbf{u}} - \mathbf{u}^o)$ is the correction vector to the approximated values; $\mathbf{B} = \frac{\partial \mathbf{F}}{\partial \mathbf{l}}$ is the design matrix of partial derivatives of observation equations with respect to the coordinates x_i and y_i ; $\mathbf{v} = (\hat{\mathbf{l}} - \mathbf{l})$ is the vector of residuals; and $\mathbf{w} = F(\mathbf{u}^o, \mathbf{l})$ is the misclosure vector. The condition equation (8) is linearised as shown in equation (11) and applied as a weighted unknown constraint:

$$\mathbf{G}_{1,u} \delta_{u,1} + \mathbf{w}_{c1,1} = \mathbf{v}_{c1,1}. \quad (11)$$

In equation (11), $\mathbf{G} = \frac{\partial \mathbf{g}}{\partial \mathbf{u}}$ is the matrix of partial derivatives of the condition equation (8) with respect to the unknowns; $\mathbf{w}_c = \mathbf{g}(n_x^o, n_y^o)$ is the misclosure vector of the condition equation (8) for the approximated values of n_x, n_y ; and \mathbf{v}_c is the residual of the constraint.

The orthogonal distance of each boundary point (i) to its corresponding line segment (j) is assumed to be the error (σ_i) and is calculated from equation (2). Accordingly, the weight matrix \mathbf{P} for the observations is designed as $\mathbf{P}_{(i)} = \text{diag}(1/\sigma_i^2)$. The weight matrix \mathbf{P}_c for the constraint must be over-weighted, for example, $\mathbf{P}_c = \mathbf{I}^2$ which has (1×1) dimensions. Otherwise, for n_x and n_y , an infinite number of solutions are obtained because the scale of the vector distances \mathbf{d} cannot be determined. Finally, the adjustment solution is formulated according to the least squares procedure considering the mixed model as follows:

$$\delta = -[\mathbf{A}^T(\mathbf{B}\mathbf{P}^{-1}\mathbf{B}^T)^{-1}\mathbf{A} + \mathbf{G}^T\mathbf{P}_c\mathbf{G}] [\mathbf{A}^T(\mathbf{B}\mathbf{P}^{-1}\mathbf{B}^T)^{-1}\mathbf{w} + \mathbf{G}^T\mathbf{P}_c\mathbf{w}_c]. \quad (12)$$

EVALUATION

Study Area

The ISPRS benchmark dataset of Vaihingen, a town in Southern Germany, includes three sites called Areas 1, 2 and 3 and were used for this evaluation. The sites contain 107 buildings with dramatically varying sizes and complexity between the three areas. The specifications of the dataset can be found in Cramer (2010).

Building Detection Results

Several evaluation approaches have been used in the literature, including object-based, area-based and RMSE (Rutzinger et al., 2009; Rottensteiner et al., 2014), together with polygon and line segments (PoLiS) (Avbelj et al., 2015). The area-based approach (using

completeness, correctness and quality) was chosen to cover all quality aspects (Potůčková and Hofman, 2016) and the results are presented in the left-hand part of Table I. The results of the object-based approach (applied twice for the two cases of buildings larger than 10 m^2 and then larger than 50 m^2), together with the method using the RMSE of the orthogonal distances between the extracted vertices and the closest vertices in the reference, are presented in the right-hand side of Table I.

TABLE I. Area-based and object-based evaluations of building outline extraction output. C_m =completeness, C_r =correctness, Q_l =quality and the RMSE in metres.

Area	Area or pixel based (%)			Object based						RMSE (m)
				>10m ²			>50m ²			
	C _m	C _r	Q _l	C _m	C _r	Q _l	C _m	C _r	Q _l	
1	92.70	96.50	89.70	93.80	100.00	93.80	100	100	100	0.79
2	95.20	94.80	90.50	91.70	91.70	90.00	100	100	100	0.73
3	87.10	97.40	85.10	87.50	100.00	87.50	97.4	100	97.4	0.73
Avg.	91.67	96.23	88.43	91.00	97.23	90.43	99.13	100	99.13	0.75

In terms of the area-based evaluation, the presented algorithm shows satisfactory results with a high average correctness of 96.23% and an average quality of 88.43% (Table I). Nevertheless, a 99.13% average quality rate has been achieved for buildings larger than 50 m^2 . The performance of the approach for the three Vaihingen sites is visualised in Fig. 6. Fig. 6(c2) depicts a building with insufficient point cloud data (highlighted by the red circles) while the interpolated DSM offered better two-dimensional (2D) grid data. This specific building was missed by numerous airborne laser scanning (ALS) based algorithms (Rottensteiner et al., 2014) due to inadequate data. Working on a gridded DSM can overcome this problem more efficiently as this building has been successfully detected (as highlighted by the black dashed circle in Fig. 6(c)).

On the other hand, there are some complex cases where the proposed algorithm struggles in its performance. For instance, Fig. 6(a1) presents a profile of part of a building missed in Area 1. It could not be detected because the height difference with respect to the terrain is less than 2 m. The other problem the algorithm faces is related to the detection of low man-made objects such as those indicated by the green arrows in Fig. 6(c) leading to false-negative errors. The main reasons for these buildings being missed are firstly, that the height difference with their surrounding terrain is quite low. For instance, the building indicated by the top green arrow in Fig. 6(c) has approximately 1 m height difference with the terrain to its left and right as shown in the profile in Fig. 6(c1). As the height difference is less than the pre-defined threshold of 2 m in Area 3, the building was not detected. Secondly, the extracted DEM has been interpolated above the terrain (Fig. 6(c4)). The DEM (blue dashed line) is slightly raised on the left, reducing the height difference to the DSM (red solid line) to approximately 1.8 m. By subtracting the DEM from the DSM to create the nDSM (mainly buildings and trees), a part of this building was missed because the height difference is less than the threshold of 2 m. By decreasing the threshold (for example, to 1.5 m), such low buildings can be detected but, at the same time, non-relevant elevated objects (such as vehicles) will be incorrectly identified as buildings. In addition, a greater area around buildings can also be detected due to the fact that the edges in the DSM are not sharp and therefore building outlines are noisier. Hence, the correctness rate might be

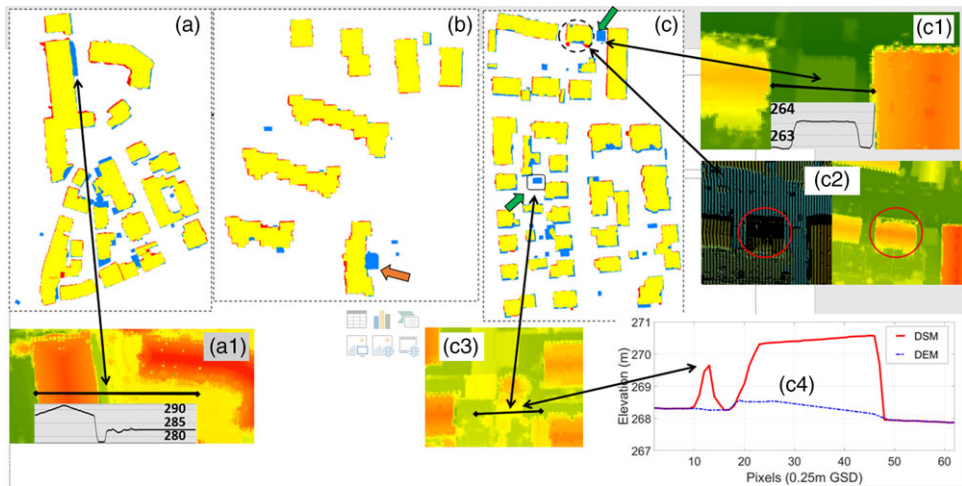


FIG. 6. Building detection results for the Vaihingen dataset: (a) Area 1; (b) Area 2; and (c) Area 3. True positives, false negatives and false positives are depicted in yellow, blue and red, respectively. Analysis of some complex scenarios are presented in (a1) and (c1) to (c4).

decreased more than the potential increase in completeness. Other disadvantages are associated with noisy edges; it may lead to an increase in the complexity of the building outline regularisation process. A similar case is indicated by the orange arrow in Fig. 6(b) where part of the building is missing because it is approximately 1 m above the DEM.

Comparison of Building Detection Performance

In comparison with the results presented in Rottensteiner et al. (2014), the proposed method shows very good performance for the area-based evaluation with an average quality of 88.43%. It is also positioned within the highest level in terms of correctness, with an average of 96.23%. Compared with the methods based on morphological filtering by (i) Mongus et al. (2014) and (ii) Zhao et al. (2016), the proposed method shows a higher average quality by 2.6% and 3.4%, respectively, over the three sites, but a lower completeness in Area 3 by 1.8% and 0.6%, respectively. In addition, the proposed method shows a significantly higher area-based quality rate for Area 1 by 4.8% and 6.2% over sources (i) and (ii), respectively. This is an indicator that morphological filters show a worse performance in sloping terrain. In comparison with the lidar data-fusion method (Du et al., 2017), the proposed method shows a lower average area-based quality by 1% but a higher average correctness by 1.36% in the three areas.

Compared to point-cloud-based approaches such as Awrangjeb et al. (2014), the proposed method performed significantly better in area-based correctness and quality (96.23% versus 91%, 88.43% versus 80.7%, respectively). Only Area 3 showed a lower completeness rate by approximately 0.6% compared to Awrangjeb et al. (2014). However, while that work assumed the DEM was available, this paper has presented a fully automatic workflow. In terms of geometrical analysis, the proposed method yielded a higher accuracy by 0.17 m (0.75 m versus 0.87 m).

Setting Parameters for Building Outline Regularisation

The goal of this section is to find reasonable ranges for the following parameters: nDSM threshold (NT), median filter (MF), RMSE threshold (RT), angle detector (AD) and similarity ratio (SR). For sensitivity analysis, five different values for each of the applied parameters were used. The following values were used for these parameters:

- (1) nDSM threshold (NT): 1.5, 2, 2.5, 3 and 3.5 m;
- (2) median filter (MF): 0, 0.5, 1, 1.5 and 2 m;
- (3) RMSE threshold (RT): 0.4, 0.8, 1.2, 1.6 and 2 m;
- (4) angle detector (AD): 1, 1.5, 2, 2.5 and 3 m; and
- (5) similarity ratio (SR): 100%, 90%, 80%, 70% and 60%.

While the nDSM threshold NT impacts the results of the building detection, it has no significant impact on the regularisation algorithm. Typical values for the nDSM threshold usually fluctuate in the range of 2 to 3 m depending on the minimum height of man-made objects in the scene. Fig. 7(a) shows that increasing the NT leads to an increase of the correctness but a decrease of the completeness. This is because small buildings and parts of buildings are lost.

In the regularisation step, a comprehensive sensitivity analysis was performed for this new approach. The RMSE of the extracted vertices was calculated as a suitable measure to estimate the planimetric accuracy of the regularisation results. In Fig. 7(b), the numbers 1 to 5 along the x axis indicate the set values for each of the five parameters ((1) to (5) above), whilst the y axis shows the calculated RMSE of the extracted vertices. When one of the parameters was changed, the other parameters were set at their standard values: the NT value depends on the lowest building that exists, $MF=1$ m, $RT=1$ m, $AD=2$ m and $SR=90\%$. Overall, RT was found to be the most sensitive parameter. Choosing a suitable RMSE threshold value is highly dependent on the complexity of buildings in the scene, and on the desirable degree of simplification. For instance, Area 2 has complex buildings with many small edges (sometimes less than 2 m) which required a strict RT value of 0.75 m (see Table II) to preserve these details. Compared to Area 3, the used RT was higher (1.1 m Table II). Therefore, some small details have been lost (Fig. 10(c)). Such details can be reinstated with a stricter RT threshold (similar to Area 2), but the regularisation goal is to obtain the highest reduction in the number of boundary points, while at the same time

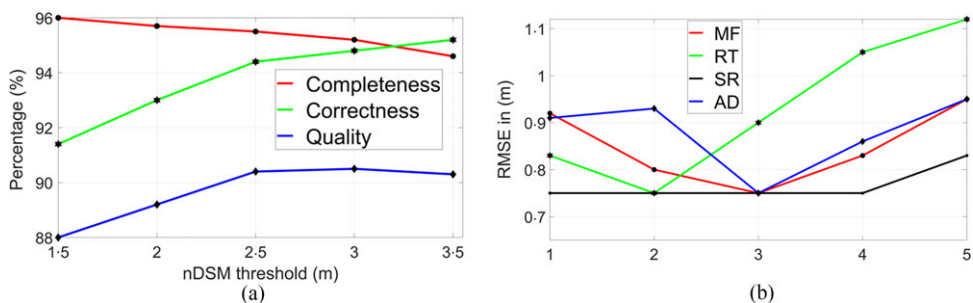


FIG. 7. Sensitivity analysis of the thresholds. (a) Sensitivity analysis of the nDSM threshold (NT). (b) Sensitivity analysis for the five parameter alternatives (1 to 5) of the MF , RMSE threshold (RT), angle detector (AD) and similarity ratio (SR).

TABLE II. Evaluation results of building reconstruction: area-based completeness (C_m), correctness (C_r), quality (Q_l), RMSE of the extracted vertices and RMSE of centroids of building objects. Parameter settings for the three sites are shown for the median filter MF , nDSM threshold NT , RMSE threshold RT , angle detector AD and similarity ratio SR .

Area	Area or pixel based (%)			RMSE of vertices (m)	RMSE of centroids (m)	Execution time (s)	Parameters set-up (m)				
	C_m	C_r	Q_l				MF	NT	RT	AD	SR (%)
1	88.7	95.6	85.3	0.93	0.8	97.69	1	2.3	1.4	1.75	75
2	93.4	95.3	89.3	0.75	0.57	67.71	1	2.7	0.75	2	~
3	84.1	96.1	81.4	1.04	0.7	94.05	0.75	2	1.1	2	68
Avg.	88.73	95.67	85.33	0.91	0.69						

maintaining the original building shape as much as possible. In the next section an extensive analysis and evaluation will be given for RT , which was used in the proposed algorithm for generating approximate building polygons or building simplification.

The MF and AD parameters show less sensitivity in behaviour and both can be fixed to their default values of 1 and 2 m, respectively (see Table II). Similarly, SR shows a more robust behaviour and only at a value of 60% does the calculated RMSE increase to 0.83 m. At 60%, the chance of a building being oversimplified and modelled as a rectangular is too high.

Evaluation of the Proposed Building Simplification Approach

In order to evaluate the performance of the proposed simplification algorithm, a comparison was performed with the DP algorithm. As a data-driven approach, the DP algorithm was chosen because it is the most common simplification algorithm in the literature, as discussed previously. In order to make both methods comparable, the parameters of both algorithms were adjusted so that the same number of vertices was obtained. As the number of vertices is identical, the standard deviation (SD) is a reasonable measure to compare the success of both approaches during the evaluation process. In fact, an algorithm that yields a lower number of vertices, associated with a lower RMSE, definitely shows a better performance.

First of all, it is highlighted that the DP algorithm is highly sensitive to the position of the starting point. Fig. 8 shows the results of the extraction of vertices depending on the starting point position (black triangles indicated by the orange arrows) for both approaches. In this figure, boundary points are shown by coloured dots assigned to their corresponding edges. The top row presents the DP algorithm results and the bottom row shows the results of the proposed method. While the DP algorithm produced significantly different output regarding the shape and number of obtained vertices (in spite of a fixed tolerance threshold (T) with value of 10 pixels (1 pixel = 0.25 m GSD)), the proposed method showed a more stable performance (the same number of vertices was obtained with a fixed value of the RMSE threshold (RT) equal to 7 pixels). Hence, where possible, it is desirable to fix the number of vertices; this is not possible in the case of the DP algorithm. The main reason is that the tolerance threshold (T) corresponds to the farthest boundary points to its assigned edges. This farthest point is not always located at a building's corner, which is the only way to detect vertices in the processing of the DP algorithm. In contrast, the proposed algorithm utilised the RMSE value as a threshold, which is more robust against outliers because these

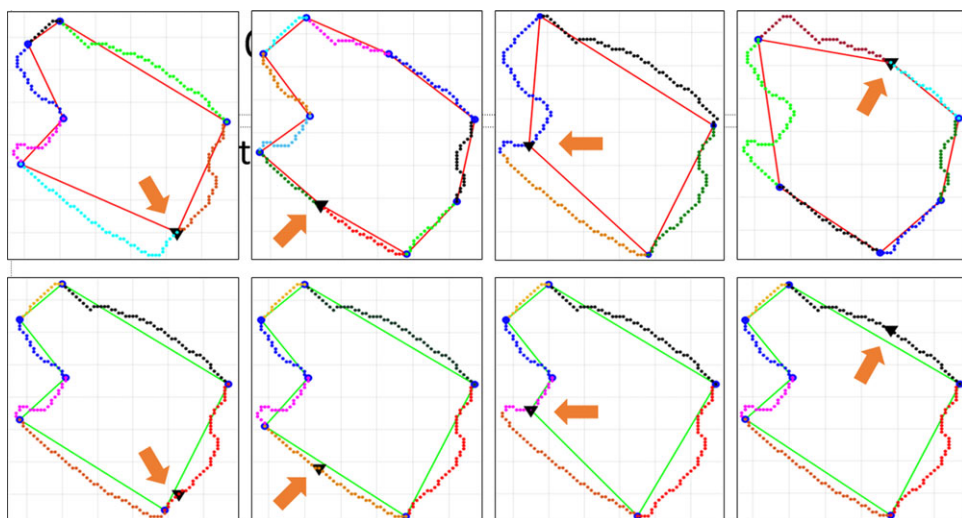


FIG. 8. The sensitivity to a change of the starting point for the DP algorithm (top row) and the proposed method (bottom row). In the DP-based algorithm, the final number of vertices changes from 6 to 9 to 4 and finally 6 (from left to right) even with one fixed tolerance threshold (10 pixels). The proposed algorithm seems more stable as the same number of vertices (6) has been obtained.

can be hidden in the calculated RMSE value. Nevertheless, corner detection has been implemented in the simplification approach of the proposed method through the likelihood equation (3).

Fig. 9 shows the results of the comparison with the DP algorithm. The figure shows eight buildings (examples) processed by both the proposed method and the DP-based method, applying different parameters for both algorithms. These parameters are: the tolerance threshold (T) for the DP algorithm and RT for the proposed approach. In addition, the calculated standard deviation (SD) values are presented in order to compare the results in an objective manner. Examples 1 and 2 show the same input building boundary points that should be represented by six vertices and six edges to achieve the best fit. However, the best possible DP simplification output is seven vertices with an SD value of 2.2. Decreasing the number of vertices to six by changing the value for T (from 16 to 19), the shape diverges from the optimal solution and produces a SD value of 3.8. In contrast, the proposed algorithm yields six vertices and a SD value of 0.96. Other performance issues are related to the ad hoc tuning of parameters of both algorithms when dealing with buildings of different complexity. For instance, simple building structures such as the third example in Fig. 9 requires a minimum T value of 8 to produce four vertices, while more complex shapes (example 6 in the figure) requires a T of at least 11. In addition, for other types of buildings (such as examples 7 and 8), it is a long iterative process to determine the T value yielding the best outcome with the DP algorithm. In contrast, the proposed algorithm requires significantly less empirical tuning of the RT value; an RT value of 4 is suitable for all types of buildings presented in Fig. 9. Overcoming such issues is essential for the automation of building outline regularisation, especially if different building complexity levels exist in one scene.

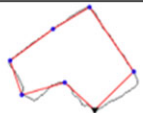
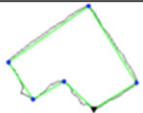
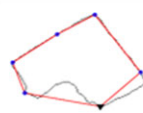
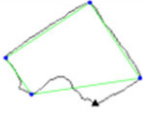
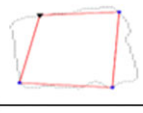
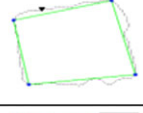
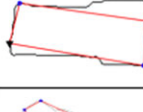
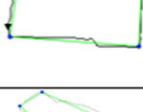
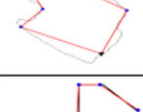
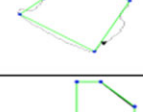
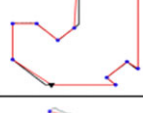
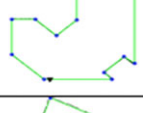
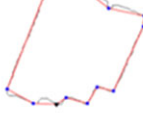
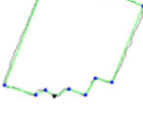


Example	DP	T (threshold)	V	SD	Proposed	RMSE (threshold)	V	SD
1		9-15	7	2.20		4-9	6	0.96
2		16-19	6	3.86		10-15	4	3.63
3		8-30	4	3.20		4-18	4	1.17
4		21-37	4	5.71		4.4-27	4	1.26
5		10	6	2.60		4.2-20	6	1.35
6		11-23	13	2.56		3-27	13	0.67
7		6	12	1.45		4	10	1.03
8		7	9	1.76		5-6	6	1.44

FIG. 9. Comparison of simplification results created by the DP algorithm and the proposed algorithm. Triangles (\blacktriangle) represent the starting boundary point. T is the tolerance threshold for DP. V is the number of vertices (\bullet), SD is the calculated standard deviation. Note the RMSE threshold is used in the proposed algorithm.

Final Regularisation Results

The final regularisation results are presented in Fig. 10. In this figure, (a), (b) and (c) represent the regularisation results of Areas 1, 2 and 3, respectively. The assessment is divided into qualitative and quantitative evaluations.

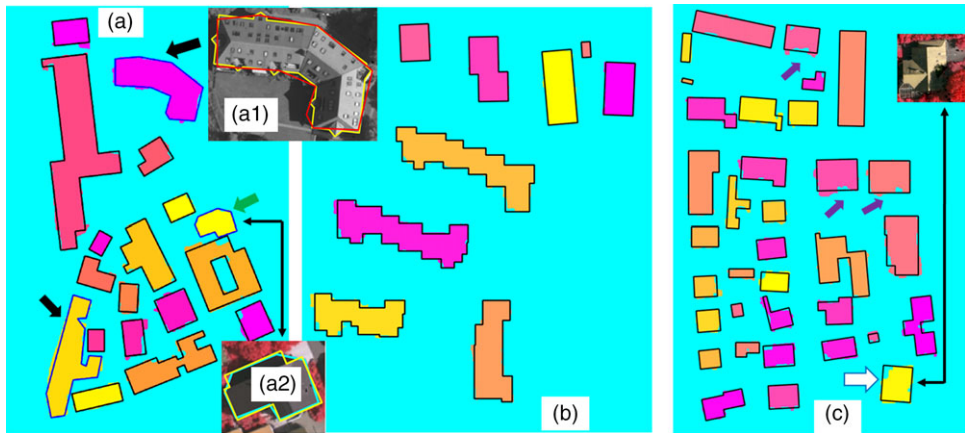


FIG. 10. Final regularisation results. (a), (b) and (c) represent the regularisation results of Areas 1, 2 and 3, respectively.

- (1) *Qualitative evaluation*: Fig. 10(a) contains complex buildings with more than two main orientation directions, which are difficult to simplify (such as buildings indicated by black arrows). Such buildings cannot be processed by a model-driven approach due to their non-rectilinearity or simplified to an overly complex polygon. Nevertheless, the results produced (red polygon) were still very close to the reference (yellow polygon) in Fig. 10(a1). While the majority of the building outlines have been successfully regularised, some rectilinear buildings, for example, the one indicated by the green arrow in Fig. 10(a), were not simplified correctly (cyan polygon) when compared with its reference (yellow polygon) in Fig. 10(a2). The main reason is that the initial solution did not satisfy the rectilinearity conditions by providing an accurate number and position of the vertices. Nevertheless, the final simplified result for this building is still close to its reference. The three buildings indicated by the purple arrows in Fig. 10(c) show a high level of noise at their edges, which makes the data-driven approach troublesome in providing an accurate solution. However, such an issue is overcome by combining this with the model-driven approach used.

On the other hand, estimating the main direction of a building is a challenging task. Overall, compared with the reference, although the final building orientation results are not identical, they are very close to the reference. The building indicated by the white arrow in Fig. 10(c) demonstrates an example of this problem. In this work, the main direction is estimated from the longest four-line segments and are introduced as an initial solution for the GH model adjustment. Then, the final dominant direction is obtained simultaneously with the solution for the best-fitting polygon edges with respect to their corresponding boundary points. However, for this specific building the problem is related to missing data caused by removing trees touching and slightly overlapping the building, as can be seen in the orthophoto in Fig. 10(c).

- (2) *Quantitative evaluation*: Building polygons are rasterised. The rasterisation process may introduce errors, but it provides a more robust outcome (Potůčková and

Hofman, 2016). Values for area-based completeness, correctness, quality, RMSE of vertices and RMSE of the centroids of the polygons, as well as the execution time, are presented in Table II. Accordingly, high rates of correctness and completeness have been achieved for all three sites with an average of 95.67% and 88.73%, respectively. The perpendicular distances between each vertex of an extracted building polygon and the nearest boundary points from the reference has been calculated. Distances greater than a pre-defined threshold of 3 m are excluded (Rottensteiner et al., 2014). The average RMSE of the geometric accuracy is 0.91 m and the RMSE of the building centroids is 0.69 m in all areas. The RMSE seems to be slightly high in all test sites. However, it is anticipated that the RMSE increases with a higher degree of generalisation and vice versa. For instance, adding more polygon vertices or edges usually leads to a lower RMSE. However, this is not the goal of the simplification process. This efficiency allows the extracted building outlines to be used for further processing steps, such as the updating of an existing building outline register.

CONCLUSION

A novel automatic workflow for building detection and boundary regularisation from 2D and 3D sensor data has been presented and evaluated using Areas 1 to 3 of the ISPRS Vaihingen benchmark dataset. Four concluding observations can be made:

- (1) The procedure for building detection has achieved a high level of area-based quality (88.43%) compared to the results presented in Rottensteiner et al. (2014). The high quality of the output for the proposed algorithm is due to its reliance predominantly on the height difference parameter for DEM extraction. It has been shown that this is the most meaningful indicator for DEM generation, as well as building detection. This means that the traditional slope parameter used for DEM generation is eliminated. Furthermore, the planarity measure successfully improves the result by removing trees in shadow areas.
- (2) An innovative data-driven approach for generating an accurate approximate building polygon is presented. In this approach the geometrical characteristics of buildings have been addressed by methodologies such as corner detection and area preservation, implemented in a likelihood function. In contrast, the majority of existing data-driven approaches utilise the DP algorithm, which does not consider the geometric aspects present in buildings. It has to be highlighted that the DP algorithm is quite sensitive to both the position of the starting point and to noise, due to the distance threshold employed whilst not considering the geometrical characteristics of buildings.
- (3) Combining the robustness of model-driven approaches with the flexibility of data-driven techniques in one comprehensive strategy proved to be efficient in the regularisation of complex building structures. Therefore, the majority of the complex buildings have been modelled accurately. However, there are some remaining difficulties, such as recovering missing data caused by trees occluding rooftops. The effect of such a problem has been mitigated significantly through the applied model-driven approach.
- (4) Regularisation of rectilinear building outlines has been largely solved. Firstly, the perpendicularity constraints of adjacent building edges must be integrated into the methodology. Secondly, the best fitting of these edges with their corresponding

boundary points has to be solved. This was achieved by minimising the orthogonal distances between them. Nevertheless, these two steps should be solved simultaneously in the adjustment process (Avbelj, 2015). Skipping one of these steps will not achieve a precise solution. In this study, both requirements are simultaneously solved by modifying the GH model adjustment.

The final regularisation results have been evaluated both qualitatively and quantitatively. Qualitatively, the proposed data-driven approach for polygonal simplification as well as model-driven approach for rectangular buildings outlining have been successfully applied by providing an accurate solution for building polygons. Quantitatively, the proposed method achieved average values of 88.7% and 95.67%, respectively, for area-based completeness and correctness. The RMSE of vertices and centroids were 0.91 and 0.69 m, respectively. The future direction of work will be testing the proposed methodology on data with lower resolution and to updating a given building map. This will be extended to applying the proposed methodology to 3D building reconstruction.

ACKNOWLEDGEMENTS

This work was supported by The Higher Committee for Education Development (HCED) in Iraq. The Vaihingen dataset was provided by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) (Cramer, 2010): <http://www.ifp.uni-stuttgart.de/dgpf/DKEP-Allg.html>. The authors would also like to acknowledge the editor and reviewers for their valuable comments.

REFERENCES

- AREFI, H., 2009. *From LIDAR Point Clouds to 3D Building Models*. Doctoral thesis, Universität Bundeswehr München, Munich, Germany. 128 pages.
- AVBELJ, J., 2015. *Fusion of Hyperspectral Images and Digital Surface Models for Urban Object Extraction*. Doctoral thesis, Technische Universität München, Munich, Germany. 141 pages.
- AVBELJ, J., MÜLLER, R. and BAMLER, R., 2015. A metric for polygon comparison and building extraction evaluation. *IEEE Geoscience and Remote Sensing Letters*, 12(1): 170–174.
- AWRANGJEB, M., 2016. Using point cloud data to identify, trace, and regularize the outlines of buildings. *International Journal of Remote Sensing*, 37(3): 551–579.
- AWRANGJEB, M., LU, G. and FRASER, C. S., 2014. Automatic building extraction from LIDAR data covering complex urban scenes. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40(3): 25–32.
- AXELSSON, P., 2000. DEM generation from laser scanner data using adaptive TIN models. *International Archives of Photogrammetry and Remote Sensing*, 33(B4): 110–117.
- BACKES, A. R. and BRUNO, O. M., 2013. Polygonal approximation of digital planar curves through vertex betweenness. *Information Sciences*, 222: 795–804.
- BRÉDIF, M., TOURNAIRE, O., VALLET, B. and CHAMPION, N., 2013. Extracting polygonal building footprints from digital surface models: a fully-automatic global optimization framework. *ISPRS Journal of Photogrammetry and Remote Sensing*, 77: 57–65.
- BULATOV, D. and LAVERY, J. E., 2010. Reconstruction and texturing of 3D urban terrain from uncalibrated monocular images using L_1 splines. *Photogrammetric Engineering & Remote Sensing*, 76(4): 439–449.
- BULATOV, D., HÄUFEL, G., MEADOW, J., POHL, M., SOLBRIG, P. and WERNERUS, P., 2014. Context-based automatic reconstruction and texturing of 3D urban terrain for quick-response tasks. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93: 157–170.
- CHEN, Q., GONG, P., BALDOCCHI, D. and XIE, G., 2007. Filtering airborne laser scanning data with morphological methods. *Photogrammetric Engineering & Remote Sensing*, 73(2): 175–185.
- CRAMER, M., 2010. The DGPF-test on digital airborne camera evaluation – overview and test design. *Photogrammetrie-Fernerkundung-Geoinformation*, 2010(2): 73–82.

- DOUGLAS, D. H. and PEUCKER, T. K., 1973. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 10(2): 112–122.
- DU, S., ZHANG, Y., ZOU, Z., XU, S., HE, X. and CHEN, S., 2017. Automatic building extraction from LiDAR data fusion of point and grid-based features. *ISPRS Journal of Photogrammetry and Remote Sensing*, 130: 294–307.
- DUTTER, M., 2007. *Generalization of Building Footprints derived from High Resolution Remote Sensing Data*. Diploma thesis, Technische Universität Wien, Vienna, Austria. 119 pages.
- FISCHLER, M. A. and BOLLES, R. C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6): 381–395.
- GERKE, M., HEIPKE, C. and STRAUB, B.-M., 2001. Building extraction from aerial imagery using a generic scene model and invariant geometric moments. *IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*, Rome, Italy. Pages 85–89.
- GONZALEZ, R. C., WOODS, R. E. and EDDINS, S. L., 2004. *Digital Image Processing Using MATLAB*. Prentice Hall, Upper Saddle River, New Jersey, USA. 827 pages.
- GROSS, H. and THOENNESSEN, U., 2006. Extraction of lines from laser point clouds. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(3/1). 6 pages.
- HARALICK, R. M., STERNBERG, S. R. and ZHUANG, X., 1987. Image analysis using mathematical morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(4): 532–550.
- HE, Y., ZHANG, C. and FRASER, C. S., 2014. An energy minimization approach to automated extraction of regular building footprints from airborne LiDAR data. *International Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2(3): 65–72.
- HÖHLE, J., 2017. Generating topographic map data from classification results. *Remote Sensing*, 9(3): article 224. 24 pages.
- HOUGH, P. V., 1962. *Method and means for recognizing complex patterns*. US Patent 3069654A.
- JWA, Y., SOHN, G., TAO, V. and CHO, W., 2008. An implicit geometric regularization of 3D building shape using airborne LiDAR data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37(B3a): 69–76.
- KILIAN, J., HAALA, N. and ENGLISH, M., 1996. Capture and evaluation of airborne laser scanner data. *International Archives of Photogrammetry and Remote Sensing*, 31(B3): 383–388.
- KRAUS, K. and PFEIFER, N., 1998. Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 53(4): 193–203.
- KWAK, E. and HABIB, A., 2014. Automatic representation and reconstruction of DBM from LiDAR data using recursive minimum bounding rectangle. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93: 171–191.
- LEE, H. S. and YOUNAN, N. H., 2003. DTM extraction of LiDAR returns via adaptive processing. *IEEE Transactions on Geoscience and Remote Sensing*, 41(9): 2063–2069.
- LIU, X., 2008. Airborne LiDAR for DEM generation: some critical issues. *Progress in Physical Geography*, 32(1): 31–49.
- LONG, J., SHELHAMER, E. and DARRELL, T., 2015. Fully convolutional networks for semantic segmentation. *IEEE Conference on Computer Vision and Pattern Recognition*, Boston, Massachusetts, USA. Pages 3431–3440.
- MAAS, H.-G. and VOSSelman, G., 1999. Two algorithms for extracting building models from raw laser altimetry data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2): 153–163.
- MARCOS, D., VOLPI, M., KELLENBERGER, B. and TUIA, D., 2018. Land cover mapping at very high resolution with rotation equivariant CNNs: towards small yet accurate models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 145(Part A): 96–107.
- MARMANIS, D., SCHINDLER, K., WEGNER, J. D., GALLIANI, S., DATCU, M. and STILLA, U., 2018. Classification with an edge: improving semantic image segmentation with boundary detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, 135: 158–172.
- MENG, X., WANG, L. and CURRIT, N., 2009a. Morphology-based building detection from airborne LIDAR data. *Photogrammetric Engineering & Remote Sensing*, 75(4): 437–442.
- MENG, X., WANG, L., SILVÁN-CÁRDENAS, J. L. and CURRIT, N., 2009b. A multi-directional ground filtering algorithm for airborne LIDAR. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(1): 117–124.
- MENG, X., CURRIT, N. and ZHAO, K., 2010. Ground filtering algorithms for airborne LiDAR data: a review of critical issues. *Remote Sensing*, 2(3): 833–860.
- MONGUS, D., LUKAČ, N. and ŽALIK, B., 2014. Ground and building extraction from LiDAR data based on differential morphological profiles and locally fitted surfaces. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93: 145–156.

- MOUSA, Y., HELMHOLZ, P. and BELTON, D., 2017. New DTM extraction approach from airborne images derived DSM. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42(1/W1): 75–82.
- PERKO, R., RAGGAM, H., GUTJAHR, K. H. and SCHARDT, M., 2015. Advanced DTM generation from very high resolution satellite stereo images. *International Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2(3/W4): 165–172.
- PIRAMANAYAGAM, S., SABER, E., SCHWARTZKOPF, W. and KOEHLER, F., 2018. Supervised classification of multisensor remotely sensed images using a deep learning framework. *Remote Sensing*, 10(9): article 1429. 25 pages.
- POHL, M., MEIDOW, J. and BULATOV, D., 2017. Simplification of polygonal chains by enforcing few distinctive edge directions. *20th Scandinavian Conference on Image Analysis. Lecture Notes in Computer Science*, 10270: 3–14.
- POTŮČKOVÁ, M. and HOFMAN, P., 2016. Comparison of quality measures for building outline extraction. *Photogrammetric Record*, 31(154): 193–209.
- ROTTENSTEINER, F., SOHN, G., GERKE, M., WEGNER, J. D., BREITKOPF, U. and JUNG, J., 2014. Results of the ISPRS benchmark on urban object detection and 3D building reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93: 256–271.
- RUTZINGER, M., ROTTENSTEINER, F. and PFEIFER, N., 2009. A comparison of evaluation techniques for building extraction from airborne laser scanning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2(1): 11–20.
- SHAN, J. and SAMPATH, A., 2005. Urban DEM generation from raw lidar data: a labeling algorithm and its performance. *Photogrammetric Engineering & Remote Sensing*, 71(2): 217–226.
- SITHOLE, G., 2001. Filtering of laser altimetry data using a slope adaptive filter. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 34(3/W4), 203–210.
- SITHOLE, G. and VOSSSELMAN, G., 2004. Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 59(1): 85–101.
- SKALLOUD, J. and LICHTI, D., 2006. Rigorous approach to bore-sight self-calibration in airborne laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 61(1): 47–59.
- SOHN, G., JWA, Y., JUNG, J. and KIM, H., 2012. An implicit regularization for 3D building rooftop modeling using airborne LIDAR data. *International Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1(3): 305–310.
- SONG, J. and MIAO, R., 2016. A novel evaluation approach for line simplification algorithms towards vector map visualization. *ISPRS International Journal of Geo-Information*, 5(12): article 223. 13 pages.
- VOSSSELMAN, G., 2000. Slope based filtering of laser altimetry data. *International Archives of Photogrammetry and Remote Sensing*, 33(B3/2: Part 3): 935–942.
- WANG, O., LODHA, S. K. and HELMBOLD, D. P., 2006. A Bayesian approach to building footprint extraction from aerial lidar data. *IEEE International Symposium on 3D Data Processing, Visualization, and Transmission*, Chapel Hill, North Carolina, USA. Pages 192–199.
- WEST, K. F., WEBB, B. N., LERSCH, J. R., POTHIER, S., TRISCARI, J. M. and IVERSON, A. E., 2004. Context-driven automated target detection in 3D data. *Automatic Target Recognition XIV. SPIE*, 5426: 133–143.
- ZHANG, K., CHEN, S.-C., WHITMAN, D., SHYU, M.-L., YAN, J. and ZHANG, C., 2003. A progressive morphological filter for removing nonground measurements from airborne LIDAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 41(4): 872–882.
- ZHANG, K. and WHITMAN, D., 2005. Comparison of three algorithms for filtering airborne lidar data. *Photogrammetric Engineering & Remote Sensing*, 71(3): 313–324.
- ZHAO, Z., DUAN, Y., ZHANG, Y. and CAO, R., 2016. Extracting buildings from and regularizing boundaries in airborne lidar data using connected operators. *International Journal of Remote Sensing*, 37(4): 889–912.

Résumé

Une méthode automatique de régularisation des contours de bâtiments est présentée, utilisant une combinaison d'approches reposant sur les données et sur les modèles pour fournir une solution robuste. La partie principale de la méthode comprend une nouvelle approche basée sur les données pour générer des polygones approximatifs à partir d'une liste de points donnés sur les contours des bâtiments. L'algorithme calcule et stocke de manière itérative des valeurs de vraisemblance entre un point initial arbitraire sur le

contour et chacun des points suivants du contour à l'aide d'une fonction issue des propriétés géométriques d'un bâtiment. L'une des étapes du prétraitement consiste à identifier les segments de bâtiments à l'aide d'un algorithme robuste de reconstruction de modèle numérique d'élévation. Les résultats d'évaluation sur un jeu de données difficile sont corrects en moyenne à 96,3% et 95,7% pour la détection et la régularisation des bâtiments, respectivement.

Zusammenfassung

Dieser Beitrag stellt eine automatisch Methode zur Regularisierung von Gebäudeumrisslinien vor, die eine Kombination von daten- und modellgetriebenen Ansätzen einsetzt, um eine robuste Lösung zu erhalten. Den Kern der neuen Methode stellt ein datengetriebener Ansatz zur Erzeugung genäherter Gebäudepolygone aus einer Liste gegebener Umrisspunkte dar. Der Algorithmus berechnet und speichert iterativ die Likelihood-Werte zwischen einem zufälligen Startpunkt des Umrisses und jedem folgenden Punkt basierend auf einer Funktion, die von geometrischen Eigenschaften von Gebäude abgeleitet wurde. In einem Vorverarbeitungsschritt müssen Gebäudesegmente mit Hilfe eines robusten Algorithmus zur Extraktion eines digitalen Höhenmodells identifiziert werden. Empirische Untersuchungen an einem anspruchsvollen Datensatz ergaben Werte für die mittlere Korrektheit von 96.3% bzw. 95.7% für die Gebäudeerkennung bzw. für die Regularisierung.

Resumen

Se presenta un método automático para la regularización de contornos de edificios, que combina aproximaciones basadas en datos y modelos para proporcionar una solución robusta. El núcleo incluye un método nuevo basado en datos para generar polígonos aproximados de edificios a partir de una lista de puntos frontera. El algoritmo calcula y almacena iterativamente la probabilidad entre un punto de frontera inicial arbitrario y cada uno de los siguientes puntos de frontera utilizando una función derivada de las propiedades geométricas del edificio. En un paso previo, los segmentos de edificio se identifican utilizando un algoritmo robusto para la extracción de un modelo de elevación digital. Los resultados de la evaluación en un conjunto de datos lograron una corrección media de 96.3% y 95.7% para la detección y regularización de edificios, respectivamente.

摘要

本文提出一种建物轮廓规则化的自动方法, 利用联合数据和模型驱动的方法来提供可靠的解决方案。该方法的核心部分包括一个新颖的数据驱动方法, 从一系列边界点生成近似的建筑物边界多边形。该算法使用建筑物的几何属性函数, 经由迭代计算并储存由任意起始边界点开始与其后各边界点位置的估值。在预处理阶段, 则须应用一可靠的算法, 由数字高程模型提取建筑物的各个组成部分。本研究以具有挑战性的数据集进行实验, 结果显示建筑物提取和规则化的平均正确率分别达到96.3%和95.7%。

GEO-LOCATING HISTORICAL SURVEY DATA AND IMAGES – A CASE STUDY FOR THE CANNING RIVER, PERTH, WESTERN AUSTRALIA

P. Helmholz¹*, Y. Mousa^{1,2}, T. Snow¹, A. Haebich¹, G. Piggott^{1,3}, J. Tonkin¹, W. Lamont¹

¹ Curtin University, GPO Box U1987, Perth WA 6845 (petra.helmholz, mousa.yousif, t.snow, a.haebich, wesley.lamont)@curtin.edu.au; (guy.piggott, jack.tonkin)@graduate.curtin.edu.au

² Civil Engineering Department, Al-Muthanna University, Al-Muthanna, Iraq

³ BHP, Level 37, 125 St Georges Terrace, Perth, WA, 6000, Australia, guy.piggott@bhp.com

Commission VI, WG IV/9

KEY WORDS: Historical maps, Western Australia, Archives, Unity, geo-visualisation

ABSTRACT:

The aim of this project is to create the required framework to allow the transformation of the Canning River location survey data captured by Dr. J. A. Ludwig Preiss from 1841 into today's maps and to utilise visualisation techniques to analyse the results. The original survey data includes distances and bearing observations as well as 14 historical maps. Firstly, the old survey data (includes distances and angles measurements) is plotted into a local coordinates system using modern surveying software (MAGNET Office). Then, common points (unchanged locations) are identified by comparison with the plotted and the current paths of the river. A similarity and affine transformation are used to find transformation parameters that allow to geo-locate the plotted river into the current geodetic datum (MGA94). The calculated Root Mean Squared Errors (RMSE) are 21.7 m and 21.1m obtained by geo-locating the common points using similarity and affine transformation, respectively. For geo-referencing the historical maps, the similarity, projective and Thin Plate Spline (TPS) transformations have been applied. It has been found that one point of interest (referred to as Nairn's house), which was drawn in one of the historical maps, still exists today (now known as Maddington Homestead). The distances from the actual position of Nairn's house to its position in the georeferenced maps using similarity, projective and TPS are 11.8m, 13m, and 14m respectively. All the gained information and map details are utilised in creating a dynamic visualisation suitable for comparing the generated map and historical map with modern aerial imaging and DTM data.

1. INTRODUCTION

Historical maps have been used and shown to be of importance, not just for historic purposes, but also for several different fields including urban planning, field cultivation and landcover/ land use (LCLU) changes (Király et al., 2008). Historical maps help to explain the evolution of areas and their dynamics over time. If the aim is to extract complex information, the precise geo-referencing of the maps is important. This paper focuses on original field observations and maps produced during a survey along the Canning River in Western Australia. They have been identified to be of great value because of the historical information of landmarks such as homesteads, significant trees, and information about the flora noted in the field sheets and on the maps.

The 1841 survey of the Canning River was commissioned by Governor Hutt of the Colony of Western Australia to record existing land grants and sites for new land settlement. Hutt appointed the German naturalist, Dr. Johann August Ludwig Preiss (1811-1883, who arrived in the colony in 1838 to lead the survey (Calaby, J. H, 1967). It was common practice to use naturalists and other men of science due to the lack of trained surveyors in the Australian colonies.

The survey was a Crown Grant Survey designed to define the actual boundaries of each grant. As per the practices of the day the original crown grant boundary posts may not have been placed by a surveyor, but the grant holder himself. Preiss may have only surveyed the line between the marks. The original Crown Grant Marks are still accepted today as being the defining point regardless of later dimensions.

The survey was completed observing the left bank of the river only, as well as the location of property boundaries. Overall, there are 2 field books available; the first field book is related to the mapping of the Canning River, while the second field book is related to the mapping of properties. Each field book contains more than 90 pages/ field sheets. The field sheets contain more information than just the survey data. This additional information is important material for historians and includes the location of homesteads, points of interests (creeks entering the river and similar) as well as botanical information recorded in field sheets.

During the 1840s a series of maps were produced based on the field sheets. Overall, there are 14 maps focusing on the location of the Canning River. The available survey data in the first field book covers nearly 2km of the river and relates to three maps. The other 11 maps available also show part of the river and other objects of interest such as property boundaries.

The State Records Office (SRO) of Western Australia has scanned the historical field sheets as well as maps in order to preserve them. These digital files were available to us as part of the project. This paper focuses only on the field sheets and maps containing information about the Canning River.

The objectives of the project are:

- To redraw the historical map based on the original field sheets
- To reference the redrawn map with the historical maps and to compare those data
- To geo-reference the re-drawn river map as well as the historical maps using Ground Control Points (GCPs) picked along the river and to assess the accuracy of the geo-referencing process

* Corresponding author

- To identify a small number of points of interest and to utilise in the accuracy assessment
- To present the information in a format in which non-spatial experts can see the process as well as the final results.

The paper is structured as follows: Section 2 provides more details about related work and Preiss' survey including the methods used, and the results of the survey. The data sources which have been used for the geo-referencing will be further introduced. Section 3 reviews the methods which have been used during the geo-referencing process and the visualisation processes and techniques and their utilisation. Section 4 presents the project results, including a detailed analysis of the results utilising the different transformation methods. Section 5 concludes the paper.

2. BACKGROUND

2.1 Related Work

Extensive research has been undertaken in the field of geo-referencing historical ortho-images or historical images. Often, the bundle adjustment using unchanged topographic features as Ground Control Points (GCP) is used in a multi-scale approach. New points are obtained from the block adjustment of the lower resolution images that can be used as control points in the block adjustment of the next scale (Zhu et al., 2008). This is an iterative process and is repeated until all resolution levels have been processed. In some cases, line features such as road network and building limits are used instead of point features (Cléry et al., 2014). In order to enhance the accuracy, sometimes additional information is used, including Digital Surface Models (DSM) derived from the historical images (Giordano, et al., 2018) or distance, orientation difference and overlap between matched lines are used. The ortho-images and images used often date no further back than the 1930s (Cléry et al., 2014), the 1940s (Giordano, et al., 2018) or even the 1960s/1980s (Zhu, et al., 2008). Giordano, et al. (2018) contains a good overview of photogrammetric and non-photogrammetric methods used in this context.

A much greater challenge is presented when attempting to geo-reference historical maps. A problem which historical ortho-images/map and topographic maps have in common, is the process of the scanning. Care must be taken when scanning the maps, i.e. using a calibrated large-format scanner, appropriate resolution, and image file format (Affek, 2013).

Nevertheless, historical satellite and airborne images are "young" compared to some historical topographic maps. Due to the age of some of these maps, the used projection and geodetic referencing systems are often unknown (Király et al., 2008).

If the used geodetic referencing system and the map projection are known, the transformation of the historical map into today's datum can be performed using the known parameters (Király et al., 2008). The challenges for this kind of transformation are well explained in (Affek, A., 2013). Furthermore, if the historical map is based on geodetic measurements (e.g. triangulation networks) and a geocentric geographic coordinate system with graticules or measured grids is used, then usually a similarity or affine transformation is used (Affek, A., 2013).

If the geodetic referencing system and the map projection are unknown, and the maps were not created based on detailed geodetic measurements but rather on estimated distances and angles, transformation parameter will have to be determined using GCPs. The selection of GCPs is usually easy and possible in topographic maps rich in information and features which do

not change over time. For instance, mountain ranges, historical roads and property boundaries. However, not all historical maps are information rich.

Topographic maps produced up to the end of the 18th century often have an unknown geodetic referencing system and map projection and have not been observed using geodetic networks. In this case a polynomial transformation and/or a rubber-sheet method are appropriate. As explained in Király et al. (2008), the rubber-sheet transformations are a group of transformations, where residual errors do not exist; the control points are transformed exactly to the to-coordinates of the points. There are some discrepancies how the transformation vectors are interpolated between the control points. The most common way is based upon the Delaunay triangulation.

Király et al. (2008) applied both transformation methods (a polynomial transformation and a rubber-sheet method) to geo-reference historical topographic maps dating back to the period from 1780-1826. The maps have been produced by the Military Survey of the Austria and have a relatively large scale of 1:12,000. The mean point error was the largest (41.9m) when using the 3rd order polynomial transformation and improved using the non-linear rubber sheeting method (38.4) and the linear rubber sheeting method (24m).

If unprocessed (raw) data is available (e.g. field sheets and notes), the challenge of different map projection and cartographic generalisation does not exist. Cartographic generalisation is the effect based on the cartographer's interpretation and adjusted to a map's scale and application (cartographic and thematic generalisation) necessary to visualise all required information on a map. Cartographic generalisation could be a source to introduce errors in the geo-referencing process. The cartographic generalisation is generally low in maps which were created out of pure cartographic interest (e.g. military maps) compared to maps which has been produced for public purposes (Balletti, 2006).

It is also important to make the geo-referenced historical maps available to the public (Previtali, 2017) and to present them in museums exhibitions in an interesting way. There are several agencies making geo-reference historical data available, e.g. projects carried out by several National Geographical Institutes. For instance, in France the IGN published the "[Carte d'État-Major](http://www.geoportail.gouv.fr/donnees/carte-de-letat-major)" (www.geoportail.gouv.fr/donnees/carte-de-letat-major-1820-1866). Another project is the [Divenire project](http://www.asmilano.it/Divenire/home.htm) by the National archive Milan (www.asmilano.it/Divenire/home.htm). After the data is geo-referenced and publicly available, the extraction of explicit map information, e.g. the location and areas of settlement is the next challenge, already addressed in several papers (Herold et al., 2011).

The resulting data, points of interest and historical information is process into a dynamic three-dimensional visualisation. Overlaying the generated map, historical map and historical points of interest with modern aerial photography and digital terrain models (DTM) will allow for the contrasting of the historic with modern imagery and reveal not only the modern equivalent of historic points of interest but also the topographical influence on the Canning River's historical variance.

2.2 Surveying methods used by Preiss

The Canning River survey was observed using a series of traverse lines with chain and bearing measurements. For distance observations links and chains were used. A chain consisted of 100 links and converts to 20.117m in today's metric system. The chain would be pulled until tense ensuring there were no kinks,

from the starting position to the end to record a distance measurement. Chain measurements were to be done at a horizontal level otherwise a correction for the slope had to be applied. Generally flat open ground was required to avoid the chain catching on obstructions.

Along with the chain, a circumferentor was used to perform direction observations. A circumferentor is a survey compass which would have been placed on top of a tripod. By lining up the target point through the eye pieces, angle readings were recorded. All angle readings in the Canning River survey were observed to the nearest minute, which was the accuracy the survey compass could be observed to.

As the Canning River survey was completed over 20km of riverbank, the task surveyor Ludwig Preiss had of keeping the survey as accurate as possible in rugged conditions is immense.

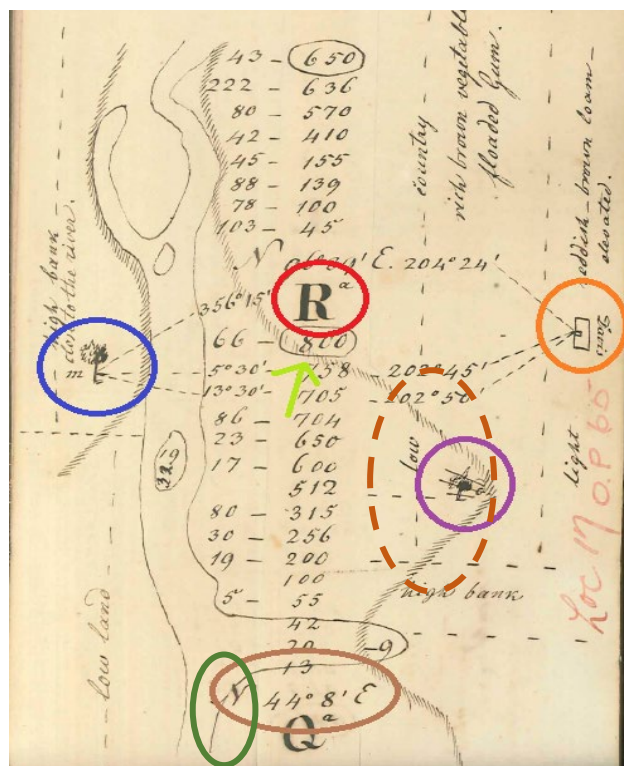


Figure 1: Field sheet of Preiss' survey of the Canning River.

It was possible to re-draw the location of the river and any points of interest using a standard surveying software (in this case MAGNET (Topcon, 2020)).

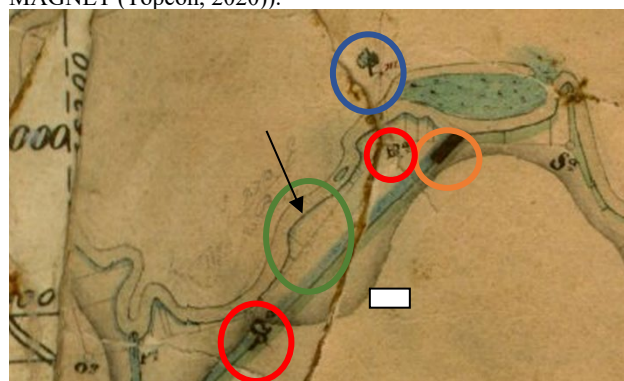


Figure 2: Section of the historical map fitting the field sheet provided in Figure 1.

A field sheet example is presented in Figure 1. The red circle shows an example of a station. Here the station is called Ra. The solid brown circle contains the azimuth from station Qa to station Ra. In this case the azimuth is $44^{\circ}8'$. Please note, that the azimuth (relative to magnetic north) and not the bearing (relative to grid/map north) are observed. The magnetic declination in this area is 5° and had been applied manually when producing the historical maps. Furthermore, the distance observations (chains) are indicated with the light green arrow in Figure 1. The offsets to the river, also in chains, are indicated with a dark green circle. If distance measurements were not possible, additional azimuth were observed enabling the determination of a location using the resection method. Such side angle readings are highlighted with the dashed brown circle. Furthermore, the orange circle is an example of a cottage located along the survey, the blue circle being a tree that was also located. The purple circle indicates a low point of the river.

The section of the river in one of the historical maps which belongs to the field sheet showing in Figure 1 is presented in Figure 2. Station Qa and Ra are both highlighted with a red circle. Furthermore, the tree which was highlighted in blue in Figure 1 is also highlighted in blue in Figure 2. The left-hand side of the river is indicated by a black arrow, and the offsets which were used to draw the river are highlighted with a green circle. While the maps are generally well preserved, there are some sections which are of poor quality due to the age of the maps. In addition, not all information has been transferred to the maps. For instance, the building highlighted in orange in Figure 1 is missing. The approximate location of this building is shown in Figure 2 (white box).

While the magnetic declination is indicated in the maps with 5° no further information about any projection is used. Therefore, also considering the overall size of the survey area, it is assumed that a planar project has been used.

2.3 Information used during geo-referencing

Two methods of geo-referencing have been identified. Firstly, some of the historical maps also contain information about property boundaries (see highlighted note in Figure 3). Those property boundaries are also shown in the map including markers for the end of a chain (red arrow in Figure 4). Those markers were also established in the field. Furthermore, the maps contain the information of the property owners as well as the approximate size of the property (blue arrow in Figure 4)

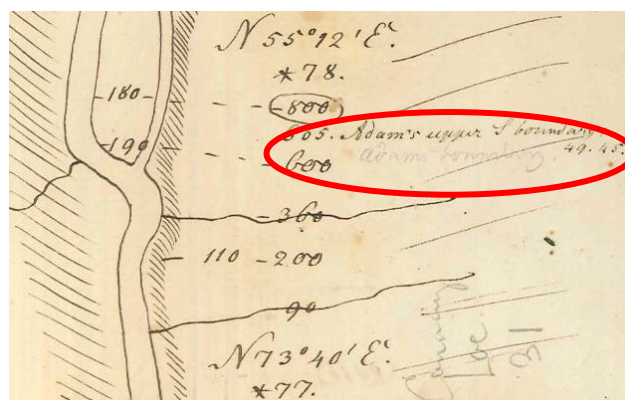


Figure 3: Field notes showing a property boundary (highlighted in red).

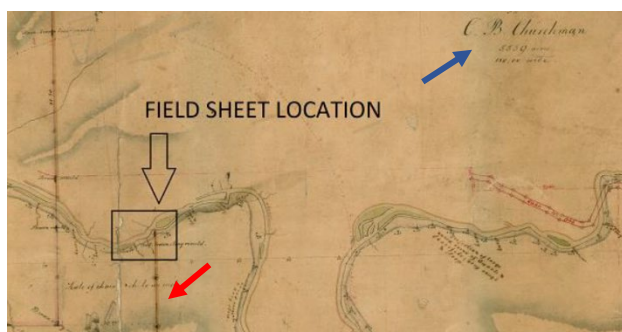


Figure 4: Historical map including property boundaries.

The most common way of changing property boundaries in Perth is through subdivisions. Therefore, the property boundaries (next to many new property boundaries) should still exist today. The attempt was made to identify the property corners, and to use those to geo-reference the historical maps. The coordinates of the property corners in Figure 5 are provided in the MGA94 coordinate system. However, careful investigation has shown that the corner points highlighted in red in Figure 5 are incorrect, and that, therefore, not always did a simple subdivision take place. Therefore, this method has been discharged.

Alternatively, the river and its topography are inspected in up-to-date airborne images and compared with the information provided in the historical maps, as well as the redrawn map based on the field sheets. It became clear that there are areas of the river which are unlikely to have changed. Therefore, it was decided to use these common points for coordinate transformations which are reviewed in the next section.

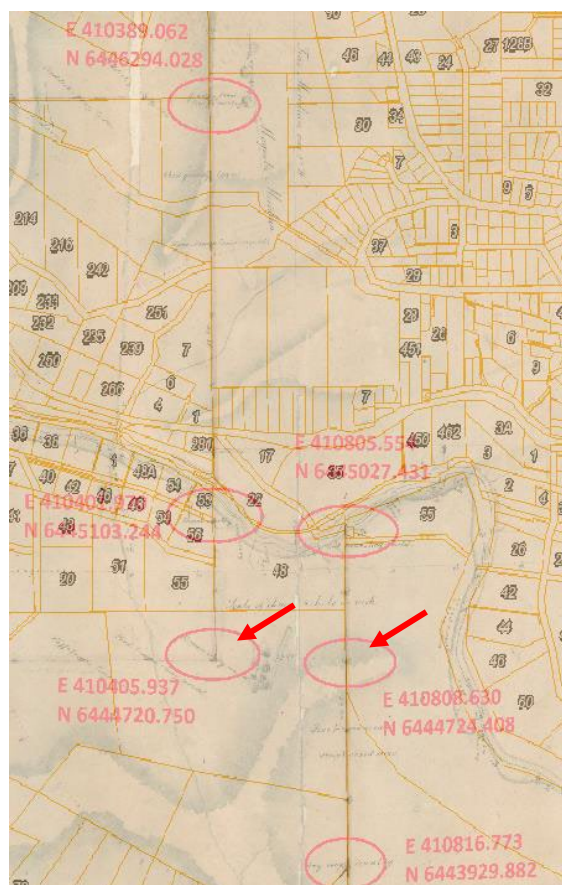


Figure 5: Historical map overlaid with today's cadastral information in the attempted to geolocate the historical map.

3. METHODOLOGY

In order to assess the accuracy of the re-drawn map as well and to perform the geo-referencing process, several different transformations have been utilised. Also, an overview of the visualisation methods and the technology used is provided.

3.1 Similarity and affine transformation

The similarity and affine transformation equations are given in (1) and (2) respectively.

$$\begin{bmatrix} X_T \\ Y_T \end{bmatrix} = \begin{bmatrix} \lambda \cos \theta & \lambda \sin \theta \\ -\lambda \sin \theta & \lambda \cos \theta \end{bmatrix} \begin{bmatrix} X_I \\ Y_I \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} X_T \\ Y_T \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} 1 & \sin \delta \\ 0 & \cos \delta \end{bmatrix} \begin{bmatrix} S_x & 0 \\ 0 & S_y \end{bmatrix} \begin{bmatrix} X_I \\ Y_I \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \quad (2)$$

In (1) and (2), X_T and Y_T are the coordinates of the target system (e.g. MGA94 system) while X_I and Y_I are the coordinates of the input system (e.g. the river local system of the re-drawing map). Both transformations account for translation parameters (Δx and Δy) as well as for a rotation (θ). While equation (1) only considers for one scale factor (λ), equation (2) considers a scale factor for x and y (S_x and S_y). Furthermore, equation (2) also considers a skew angle (δ). The parameters can be determined using common points in a least square adjustment. Afterwards, the parameters can be applied to all other points in order to perform the transformation of these points from the input system into the target system. For the process of finding the transformation parameters as well as for the transformation of points a self-implemented MATLAB code was utilised.

3.2 Projective transformation

Projective transformation is a non-linear transformation method that implicitly estimates the scale, rotation and tilt of coordinate system planes. While the similarity transformation solves for 4 parameters, and the affine transformation solves for 6 parameters, the projective transformation solves for 8 ($a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2$) parameters using the equations (3) and (4).

$$x_T = \frac{a_1 x_I + a_2 y_I + a_3}{c_1 x_I + c_2 y_I + 1} \quad (3)$$

$$y_T = \frac{b_1 x_I + b_2 y_I + b_3}{c_1 x_I + c_2 y_I + 1} \quad (4)$$

For the project transformation the implementation in QGIS (version 3.2) (Lennert, 2017) has been utilised.

3.3 Thin Plate Spline (TPS) transformation

Referring to a thin sheet of metal, this transformation has its physical analogy involving bending of the metal sheet. Hence, it is not a ridged transformation like the transformation introduced previously. TPS enables local deformations in the data, implying a penalty involving the smoothness of the fitted surface. Therefore, this spline-based technique for data interpolation and smoothing is especially useful when very low-quality originals are being georeferenced. The residuals for all used control points are nearly zero, as the thin plate is made fitting through those points. For further details see Duchon (1976).

For the TPS transformation the implementation in QGIS (version 3.2) (Lennert, 2017) has been utilised.

3.4 Visualisation technique and technology

Firstly, the re-drawn map, historical map and aerial photography are combined with a digital elevation model (DEM) of the area derived from LiDAR at a 5m resolution from 2015 (GeoScience

Australia, 2020). These maps are then overlaid within the Unity application (Unity Technologies, 2020) to develop the visualisation. Additional functionality was implemented and included:

- Adding and changing the height scale of the DEM data dynamically with different maps
- Changing the different maps' visibility dynamically
- Automating a camera viewing path over areas
- Water, Texture and Lighting adjustments to enhance realism

Then, a second visualisation was also implemented which took the three maps and DEM data and was used to create a point cloud of each map overlaid on the DEM data. This was achieved utilising Maya 2019 (Autodesk) to create a mesh object textured with each map which was then converted to a point cloud using Present (Euclidean Holographics). In addition, Open Street Map data (also projected using MGA94) has also been used.

The visualisations are presented in the Curtin University HIVE facility with three different display screen technologies.

Firstly, a high-resolution Tiled Display is used to view the visualisation at a screen resolution of 7860 x 3240 pixels (24 MPix) with a screen area of 10.028m². The Tiled Display allows for the most visually detailed image to allow direct comparison between the different maps to look for variations and some of the historical hand drawn details due to the possibility to present fine detail clarity on the map data.

Secondly, a curved Cylinder Display with stereoscopic projectors with time-sequential shutter glasses was used. This is 5048 x 1200 pixels resolution with a screen area of 37.699m². The Cylinder Display provided stereoscopic rendering on a very large surface. This provided the most accessible perception of the terrain surface comparatively to the mapped data. Furthermore, the Cylinder Screen adds stereoscopic with time sequential shutter glasses, which allows three-dimensional perception of the topographical data with the different maps aligned for comparison.

Thirdly, there was the Holographic Table (Euclidean Holographics) with four stereoscopic projectors underneath and IR tracked time-sequential shutter glasses for two users with stereoscopic views of the point cloud data with corrected projection. It allows the two users with separate 3D glasses to view different views of the terrain which are adjusted to their viewing point to give highly accurate representation of the map data in stereoscopic 3D. The table has a resolution of 1440 x 1440 pixels and a screen area of 1.6129m².

4. RESULTS AND DISCUSSION

In this section, firstly the transformation of the re-drawn map into today's geodetic datum is performed. Then follows discussion of how redrawing maps based on the original field sheets compares with the scanned historical maps. The section closes with the visualisation of the data.

4.1 Transformation of the re-drawn map based on the original field sheets to today's geodetic datum

Firstly, the previously mentioned magnetic declination of 5 degrees has been applied to the re-drawn map. Next, common points of suspected unchanged locations are identified by comparison the re-drawn map and the current paths of the river in OSM. Overall, 17 common points have been identified and utilised in a similarity (equation 1) and affine transformation (equation 2) in order to define the transformation parameters.

As the redrawn map is in a local system with a map scale factor of one, and considering the map scale factor of 0.9997156 for the MGA system in which the today's maps are defined, it is assumed that a scale factor of this dimension can be found. Furthermore, it is assumed that there is no large rotation as the magnetic declination has been applied already. There is no meaning of the magnitude to the translation vector as the first point of the re-drawn map was given arbitrary coordinates which, by purpose, have been selected to be close to the true location.

The results of the affine and similarity transformation are presented in Table 1.

Parameters	Similarity	Affine
Rotation (Minutes)	27	22
Scale	0.999949	~
Δx (m)	27.0611	27.0611
Δy (m)	-65.7418	-65.7418
S_x	~	0.9934
S_y	~	1.0018
Skew (Minutes)	~	-16

Table 1: The similarity and affine transformation parameters for the geo-location of the re-drawn map based on the historical field sheets.

While the magnitude of the translation is not important, the translation parameters are similar for both translations. The scale parameters of both transformations are also similar and similar to the predicted scale factor of 0.9997156.

The calculated rotation angle using the affine transformation is slightly lower than the rotation angle calculated using the similarity. The difference is 5 minutes and the magnitude of the rotation is around 25 minutes, and therefore below the tolerance of the used compass for the angle readings. The calculated Root Mean Squared Errors (RMSE) of the 17 common points are 21.7m and 21.1m for similarity and affine respectively are high.

Next, all other survey points (1376 points) including stations, river boundary and points of interest are transformed from the local system of the re-drawing map to the MGA94 using the calculated similarity and affine transformation parameters. The RMSE of the 1376 points geo-located by similarity and affine parameters are 7.6 m in X and 3.5 m in Y between them. These differences are insignificant considering the accuracy of the used common points, and the fact that river points may have changed in the last 180 years. Thus, both methods are sufficient to obtain accurate geo-located coordinates. Figure 6 shows a small section of the river points from the re-drawing map (black dots) overlaid on OSM. Some of the differences are very large, and maybe can be explained with an actual change of river points.

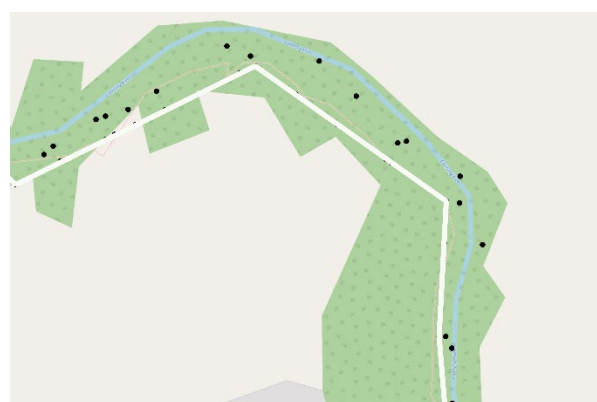


Figure 6: Section of the geo-located river points (black dots) on OSM.

As a conclusion, the absolute accuracy (positional accuracy) of the re-drawn map to today's geodetic datum is in the magnitude of less than 10m.

4.2 Transformation of historical maps to geo-referenced re-drawn map

Next, the aim is to compare the relative accuracy of the re-drawn map to the historical maps. Affek, A. (2013) suggest that if the historical maps are based on geodetic measurements (e.g. triangulation networks) and a geocentric geographic coordinate system with graticules or measured grids is used, then usually a similarity of affine transformation is applied. While both are not true for the re-drawn map and the historical map, both should be in the same coordinate system. Therefore, for the transformation of those datasets, the similarity transformation is applied. However, instead of applying the affine transformation the projective transformation is used to account for any errors which maybe have been introduced during the scanning process.

Point-labels of the re-drawn map and the historical maps are identical, which made it easy to define common points. Overall, for each historical map 10 common points have been identified and utilised a similarity transformation (equation 1) and a projective transformation (equation 2).

Due to problems with the scanning process the maps are not to scale. Furthermore, the scanned images have a pixel coordinate system and the re-drawn geo-located maps is in the metric system. Hence, large scaling numbers are suspected. Furthermore, the historical maps have not been scanned with north to the top of the page; so the rotation angles also do not provide with any further information. As none of the transformation parameters provide any valuable information, the focus next is only on the residual plots.

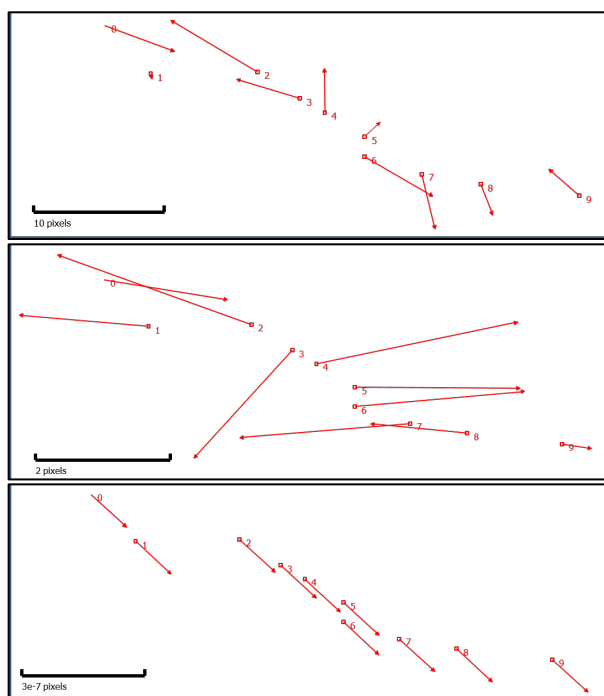


Figure 7: Residual plot for the transformation of the historical map using a similarity transformation (top), projective transformation (middle) and Thin Plate Spline (TPS) transformation (bottom). 1 pixel is approximately 1m.

The residual plots for the similarity transformation is provided in Figure 7 (top), and for the projective transformation in Figure 7

(middle). Please note that the provided scales for the plot are different. Overall, the magnitude of the residuals of the projective transformation is much smaller. An important aspect is that the residuals are all small, similar and random. Therefore, no systematic trend has been detected.

The magnitudes of the residuals are also clearly visible in the calculated RMSE values in Table 2. The RMSE of the projective transformation is half of the RMSE of the similarity transformation.

	RMSE in X (m)	RMSE in Y (m)	RMSE in XY (m)
Similarity	3.467	2.529	4.292
Projective	2.174	0.649	2.268

Table 2: RMSE for the similarity the projective transformation.

As an additional transformation, the Thin Plate Spline (TPS) transformation is applied. As this transformation creates zero residuals the residual plot in Figure 7 (bottom) is as expected. Please note that the scale bar shows a very small-scale number in this figure.

Overall, it can be concluded that the field notes (re-drawn map) and the historical map fit well which is what has been expected. The relative accuracy is within 5m.

4.3 Assessment of the positional accuracy

The final evaluation step is to evaluate the positional accuracy of the transformed historical maps. For this assessment, an object in the historical maps is located which still (or at least its ruins) exist today. This object is the Nairn House. The house burnt down in the 1930s and therefore only the skeleton of the buildings can be seen today. No other objects shown in the historical maps and field sheets could be identified as still existing today.

Nevertheless, the distances from right corner of Nairn House to its corresponding location in the geo-referenced historical maps created using the similarity, projective and TPS transformation are shown in Figure 8. Today's location of the Nairn corner has been determined using high resolution airborne images provided by EagleView and is also shown in Figure 8.

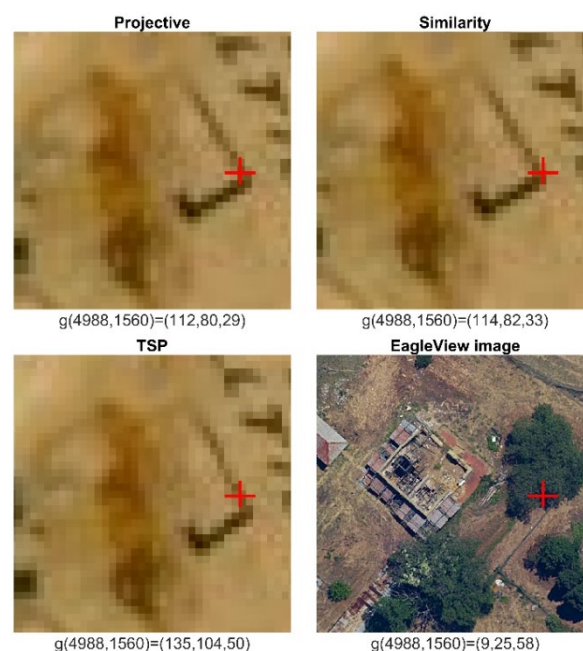


Figure 8: Positional accuracy assessment of the historical maps using Nairn House.

The distance of the corner of the similarity, projective and TPS transformation map to today are 11.8m, 13m and 14m respectively. These are very good results and expected based on previous conclusions about the absolute and relative accuracy assessments of approximately 10m and 5m respectively. It is interesting to note that the similarity transformation produced the best results for this test. This transformation has created the largest residuals in the previous relative and absolute accuracy assessment. Possible explanations are the overfitting of the projective and TSP transformation. However, it must be also pointed out that the historical map was slightly damaged close to the Nairn House which is also visible in Figure 8. This can be another contributing factor for the similarity transformation producing the best result.

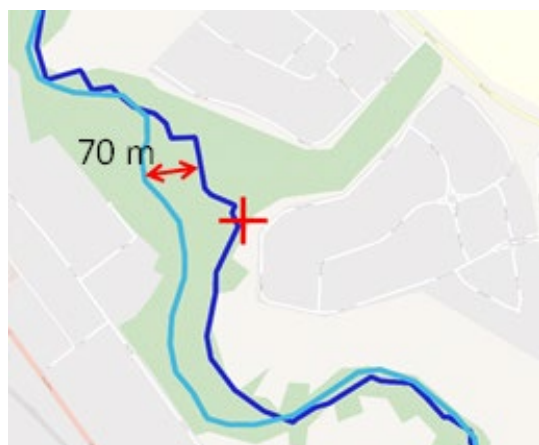


Figure 9: Area of significant change of the Canning River. In dark blue the location of the river in 1841; in light blue the location of the river today (OSM).

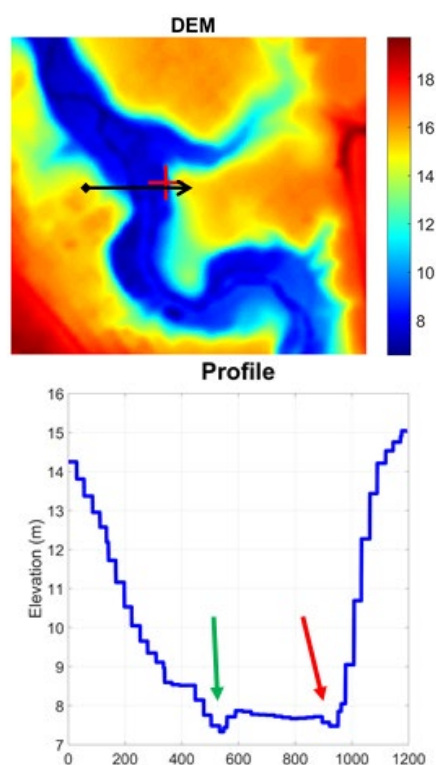


Figure 10: A profile through the DEM for the section indicated by the black arrow is presented in the top figure. In the profile on the bottom: the red and green arrows indicate the old and the current locations respectively of the Canning river.

From the data it is further concluded that there are some significant changes in the watercourse of the Canning River between 1841 and today. Figure 9 shows that the river moved approximately 70m as indicated by the red double headed arrow. However, the top and bottom parts in this figure show a strong correspondence with no significant changes.

The Digital elevation model (DEM) of the area shown in Figure 9 is presented Figure 10. A profile for the section indicated by the black arrow in the top figure is presented in the bottom figure. In the bottom profile, the red and green arrows indicate the old and the current locations of the Canning river. This is clearly indicating the movement of the river to left with nearly 70 m. The direction of the watercourse is from the bottom to the top.

Finally, Figure 11 shows the location of the historical maps with EagleView's high resolution images shown in the background, and the location of river taken out of the re-drawn map. Overall, the alignment and accuracy which was calculated for the Nairn House can be confirmed using the overlay.



Figure 11: High resolution airborne image with geo-located historical maps.

4.4 Visualisation

Visualising the project data has produced several benefits. Firstly, an accessible real time visualisation that is most suitable for public display of the generated and historic maps. Secondly, overlaying the historic content with modern airborne images (MGA94) (provided by EagleView) allows the results to be viewed and interrogated in real time alternating between the modern historical images of the same point. Thirdly, adding the Digital Elevation Data (provided by GeoScience Australia) allows the ability to visualise the historic imagery on top of the historic data and reveal the river path topographically as shown in Figure 2.6 (top figure).

This visualisation shown in Figure 12 (bottom figure) provided a visual and dynamic method to interrogate the data and visualise the conclusions outlined within this paper.

Other benefits:

- The correlation of the historical map with the re-drawn map
- The correlation of the historical map and the re-drawn map to the modern Aerial photography
- The correlation of the historical map to the topology of the terrain for the path of the river
- The correlation of buildings within the modern Aerial photography with historical referenced buildings and the discovery of the Nairn House association.

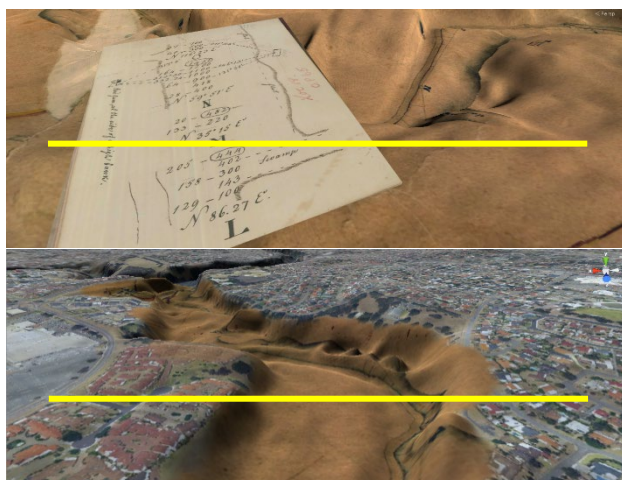


Figure 12: Top: Historical Map visualised with DEM in Unity with associated field sheet. Bottom: Modern aerial photography and historical map combined with DEM in Unity visualisation. The yellow bar indicates the same profile.

The visualisation is built as a custom application which allows for the ease of demonstration and access for the public, collaborators, and other researchers.

5. CONCLUSION

In this work, a framework for transforming old survey measurements (bearing and distances) as well as historical maps into today's geodetic referencing system (here: MGA 94) has been introduced. The paper investigated the absolute (app. 12m) and relative accuracy (app. 5m) utilising a number of different transformations. The predicted accuracies could be confirmed utilising the last shown in the 1841 maps and field sheets which still exist today. Furthermore, a pathway for further processing has been outlined in this paper that was implemented to allow non-experts to view the produced results in an accessible visual format.

The results presented here can now be used to reference the field sheets where Preiss drew and named botanical information about the river and to begin mapping the plant biodiversity he recorded. Much more can be done using geo-spatial visualisation to engage publics in new conversations about the river environments. The Preiss field books and maps have a unique role to play. As well as being 180-year-old documents of early land grants to colonists along the Canning River they are also visually compelling images of the millennia-old riverscape and of the abundance of freshwater lakes and fertile wetlands and dense stands of trees and plants. Preiss documented this moment before the staggering environmental transformations of the river began. With several river projects starting in Perth and a new state museum this is an opportune time for the project to show the way using geo-spatial mapping and visualisation.

ACKNOWLEDGEMENTS

We would like to acknowledge Anna Haebich for introducing Preiss and his maps and field books to the project. Furthermore, we would like to thank EagleView for making the high-resolution airborne images available for this project; the State Record office for their support of tracing historical images and information; and the Curtin CIC and HIVE for supporting the project by providing a scholarship.

REFERENCES

- Affek, A., 2013. Georeferencing of historical maps using GIS, as exemplified by the Austrian Military Surveys of Galicia. *Geographia Polonica*. Volume 86, Issue 4, pp. 375–390. <http://dx.doi.org/10.7163/GPol.2013.30>
- Balletti, C., 2006. Georeference in the analysis of the geometric content of early maps. *e-Perimetreon*, Vol.1, No. 1, Winter 2006, 32-42.
- Cléry, I., Pierrot-Deseilligny, M., Vallet, B., 2014. Automatic georeferencing of a heritage of old analog aerial photographs. *Int. Annals. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-3, 33- 40. doi:10.5194/isprsannals-II-3-33-2014
- GRASS Development Team, 2017. Geographic Resources Analysis Support System (GRASS) Software. Open Source Geospatial Foundation. grass.osgeo.org (20 September 2017).
- GeoScience Australia, 2020. Digital Elevation Model (DEM) of Australia derived from LiDAR 5 Metre Grid. <https://ecat.ga.gov.au/geonetwork/srv/eng/catalog.search#/meta/data/89644>. Last accessed 29/01/2020.
- Giordano, S., Le Bris, A., Mallet, C., 2018. Toward automatic georeferencing of archival aerial photogrammetric surveys. *Int. Annals. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-2, 105 – 112. <https://doi.org/10.5194/isprs-annals-IV-2-105-2018>
- Herold, H., Roehm, P., Hecht, R., Meinel, G., 2011. Automatically georeferenced maps as a source for high resolution urban growth analyses. *Proceedings of the 25th ICA International Cartographic Conference*, 5 pages.
- Király, G., Walz, U., Podobnikar, T., Czimmer, K., Neubert, M., Kokalj, Ž., 2008. Georeferencing of historical maps – methods and experiences. *Spatial Information Systems for Transnational Environmental Management of Protected Areas and Regions in the Central European Space*, Publisher: Rhombos-Verlag, Editors: Csaplovics, Elmar and Wagenknecht, Stefan and Seiler, Ulrike, 53-63.
- Lennert, M., GRASS Development Team, 2017. Addon i.segment.stats. Geographic Resources Analysis Support System (GRASS) Software, Version 7.2, Open Source Geospatial Foundation. grass.osgeo.org/grass7/manuals/addons/i.segm.ent.stats (1 June 2017).
- Previtali, M., 2017: GEOPAN AT@S: A brokering based gateway to georeferenced historical maps for risk analysis. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2/W5, 583-589. doi:10.5194/isprs-archives-XLII-2-W5-583-2017.
- Topcon, 2020. MAGNET Software Suite - Positioning software that puts you in control. <https://www.topconpositioning.com/gb/magnet-software-suite> Last accessed 28/01/2020).
- Unity Technologies, 2020. <https://unity.com/> Last accessed 29/01/2020
- Zhu, L., Erving, A., Koistinen, K., Nuikka, M., Junnilainen, H., Heiska, N., Haggren, H., 2008. Georeferencing multi-temporal and multi-scale imagery in photogrammetry. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXVII. Part B5, 225 – 230.

NEW DTM EXTRACTION APPROACH FROM AIRBORNE IMAGES DERIVED DSM

Yousif Abdul-kadhim Mousa^{a,b*}, Petra Helmholz^a, David Belton^a

^a Department of Spatial Sciences, Curtin University, Perth, WA, Australia –

^a y.mousa@postgrad.curtin.edu.au, petra.helmholz@curtin.edu.au, d.belton@curtin.edu.au

^b Al-Muthanna University, College of Engineering, Civil Engineering department, Al-Muthanna, Iraq

Commission II, WG II / 4

KEY WORDS: Digital Surface Model (DSM), Digital Terrain Model (DTM), DTM extraction, normalised DSM (nDSM), airborne images, LiDAR

ABSTRACT:

In this work, a new filtering approach is proposed for a fully automatic Digital Terrain Model (DTM) extraction from very high resolution airborne images derived Digital Surface Models (DSMs). Our approach represents an enhancement of the existing DTM extraction algorithm *Multi-directional and Slope Dependent (MSD)* by proposing parameters that are more reliable for the selection of ground pixels and the pixelwise classification. To achieve this, four main steps are implemented: Firstly, 8 well-distributed scanlines are used to search for minima as a ground point within a pre-defined filtering window size. These selected ground points are stored with their positions on a 2D surface to create a network of ground points. Then, an initial DTM is created using an interpolation method to fill the gaps in the 2D surface. Afterwards, a pixel to pixel comparison between the initial DTM and the original DSM is performed utilising pixelwise classification of ground and non-ground pixels by applying a vertical height threshold. Finally, the pixels classified as non-ground are removed and the remaining holes are filled. The approach is evaluated using the Vaihingen benchmark dataset provided by the ISPRS working group III / 4. The evaluation includes the comparison of our approach, denoted as Network of Ground Points (NGPs) algorithm, with the DTM created based on MSD as well as a reference DTM generated from LiDAR data. The results show that our proposed approach over performs the MSD approach.

1. INTRODUCTION

Having an accurate and reliable DTM is beneficial for numerous mapping applications in photogrammetry and remote sensing, such as object detection. High resolution stereo images from airborne or satellite platforms can achieve sub-meter Ground Sample Distance (GSD) and therefore have yielded the opportunity to produce a high resolution and accurate Digital Surface Models (DSMs) by using dense image matching technique (Hirschmuller, 2008). In the context of this paper a DSM is defined including all visible ground details, i.e. the visible terrain and all objects such as buildings and trees on the terrain. Therefore, a DSM can be separated into a Digital Terrain Model (DTM) representing the bare ground including roads and low vegetation, as well as a normalized DSM (nDSM) describing non-ground objects such as buildings and vegetation. However, the extraction of reliable DTMs from DSMs is not a straight forward process, and is an ongoing research topic especially with respect to densely built-up areas (Krauß et al., 2011).

The two most common approaches to generating DSMs are based on images using stereo image matching techniques and Light Detection and Ranging (LiDAR). Recently, state-of-the-art dense image matching approaches such as Semiglobal Matching SGM (Hirschmuller, 2008) have been considered to generate high resolution and accurate DSM for object detection and 3D reconstruction (Bulatov et al., 2014). However, DSM derived from stereo image matching often contains holes as a result of occlusion and mismatches (Krauß et al., 2015). Such holes can be filled by interpolation (Krauß & d'Angelo, 2011). As a result, sharp edges e.g. building boundaries might be smoothed. In contrast, LiDAR data yields more well defined DSMs and the objects outlines are well defined (Perko et al., 2015; Tian et al., 2014).

This research proposes a robust DTM extraction algorithm for DSMs derived from very high resolution airborne images in structurally complex regions. The basic idea is to enhance the approach of (Perko et al., 2015) by adding additional parameters for the selection of the minimum ground points and the pixelwise classification. Furthermore, instead of applying the complex and non optimal local slope correction of (Perko et al., 2015), an alternative technique is proposed. This technique is

simpler, more reliable and faster, which is based on creating a network of ground points (NGPs). In addition, our approach will not be affected by the smoothed transition caused by occlusion during the generated DSM process as the slope angle threshold is eliminated from the pixelwise classification process completely.

This paper is structured as followed. The related studies are reviewed and discussed in the next section. Then, we explained our approach in detail in the third section. In the fourth section, the approach is evaluated; results are shown and analysed. The last section concludes the paper and outlines future work.

2. PREVIOUS WORK

Based on the literature, there are a limited DTM extraction methods in the context of photogrammetric DSMs when compared with LiDAR-based DTM extraction methods (Beumier & Idrissa, 2016). A good review about DTM extraction algorithms for LiDAR data can be found in (Meng et al., 2010). More generally, the authors classified ground filtering algorithms into six major categories including Segmentation and Clustering, Morphological, Directional Scanning, Contour, Triangulated Irregular Network (TIN), and Interpolation. In contrast, our work targets DTM extraction from DSMs based on photogrammetry. However, while DSMs derived from airborne or very high satellite stereo images are in general similar to DSMs derived from LiDAR, the main exception is the spatial resolution (point density). DSMs derived from images generally have a higher spatial resolution but less accuracy in height measurements (Beumier & Idrissa, 2016). Fusion of images information with DSMs could be a useful option for DTM generation. However, the variety of the man-made objects and their occurrence mitigates the anticipated benefits (Beumier & Idrissa, 2016). For these reasons our literature review is limited to approaches designed for automatic DTM extraction from only photogrammetry-based DSMs.

A *Morphological Filters* approach is proposed by Haralick et al. (1987) and Förstner (1982) and is based on the idea that -DSMs can be represented as a grayscale image with its pixel values indicating a height value. Treating the DSM as a grayscale image gives the opportunity to apply image processing technologies which can remove high (bright) areas from the

DSM. Namely, an opening filter consisting of erosion and dilation are used to eliminate non-ground points (e.g. buildings) from the DSM. For instance, (Krauß et al., 2011) first applies a minimum filtering with an approximated filter size higher than the cross section of a building. In this context, the filter size (diameter) is called a structural element (SE). As a result, all pixels containing non-ground information, e.g. roof pixels are replaced by minimum ground elevation heights within the SE. Next, a maximum filtering (dilation) is applied to restore edges of eroded terrain points. The main disadvantage of this method is the failure for DSMs containing roof objects smaller or larger than the implemented SE. Additionally, applying only classical opening on noisy DSMs containing negative outliers values leads to dominate these negative values in the resulting DTM (Krauß et al., 2011). Krauß et al. (2008) overcomes the noisy DSMs problem by applying low and high rank median filters instead of the erosion and dilation filters respectively. However, the decision of the rank of the low pass filter (e.g. 3%, 4% or 5%) depends on the applied filter size SE and the density of the built-up areas within the area of interested. Therefore, manual iterative parameter estimation is required. Furthermore, this low-rank percentage might correspond to non-ground regions, especially in high density built up areas e.g. on top of buildings, leading to dominate non-ground values in the resulting DTM. Or, vice versa, this low percentage might belong to too low bare-ground regions leading to dominate too low values.

Arefi et al. (2009) proposed a DTM extraction method named *Geodesic Dilation* which applies a vertical height threshold instead of the horizontal opening threshold. Again, as in the previous approach, the grey values correspond to the elevation heights. Two equal size images are required called mask (J) and marker image (I). The marker image grey values must be less than or equal to the mask image grey values. The marker image is generated according to:

$$J(i,j) = \begin{cases} I(i,j), & \text{if pixel } p(i,j) \text{ locates on the border of mask } (I) \\ \min(I), & \text{otherwise} \end{cases}$$

That means that the marker image has the same elevation values as the mask on its border and all other marker's pixels have one value correspond to minimum value from the mask. At the beginning, the mask image has the same value as the original DSM. For each pixel in the evaluation process, 4 directional filters are conducted from each corner to the opposite site of the image, i.e. from upper left (UL) corner to lower right (LR), from LR to UL, from upper right (UR) to lower left (LL) and from LL to UR. For each evaluation process, three pixels values from the marker are compared with three pixels values from the mask along the scanline direction, and the height difference between them is calculated. A scanline denotes a one directional line where there are a number of pixels are positioned along the line within a specific window size in the raster DSM. Non-ground objects are identified if the height difference is larger than a pre-defined threshold. However, when buildings are positioning close to the border of the image, the height difference is often lower than the threshold depending on the object surface properties. This can create non-satisfactory filtering result, especially in high resolution DSMs. Furthermore, the height difference between pixels belonging to raised bare-ground regions form the mask with their connected pixels from the marker is might be higher than the threshold, especially in sloped areas. As a result, these raised ground pixels will be eliminated from the resulting DTM.

Krauß and Reinartz (2010) proposed a DTM extraction method called *Steep Edge Detection*. The idea is to apply two median filters with different filter sizes. The different filter sizes will show the occurrences of various sharp ends. For instance, a larger dimension median filter fills up small holes while a smaller median filter tracks the elevation structure of the original DSM more precisely. Afterwards, the median filter results are subtracted from each other using a vertical threshold set to the lowest values for possible sharp edges. The resulting regions normally correspond to the bare-ground regions which

are then filled and interpolated to create the DTM. The main drawback of this method is that when some large objects are located on the building roofs, these objects maybe confused with buildings instead of being identified as part of the buildings leading to the lower roof pixels being incorrectly detected as ground pixels (Krauß et al., 2011). Secondly, when low vegetation such as small bushes are located in close proximity to buildings, these bushes might be also taken as ground points leading to decrease the accuracy of the generated DTM.

Perko et al. (2015) developed a new DTM extraction algorithm for DSMs derived from very high resolution satellite images called *Multi-directional and Slope Dependent* (MSD). This algorithm is an extension of the directional filtering approach introduced by (Meng et al., 2009). The main idea of the MSD algorithm is to specify points in the derived DSM which are positioned on bare ground regions and eliminate all other non-ground points. First of all, a robust slope fitting is performed using 2D Gaussian smoothing filter in order to smooth and correct the local slope terrain. Then, points located on bare ground regions are determined by applying 4 directional scanlines which intersect at the pixel under evaluation in the middle of a pre-defined window (Figure 2a). The dimension of the window is defined by the filtering size. For each scanline in this window, the pixel value with the minimum elevation value is selected. During the evaluation process, the elevation of the point under examination is compared to this minimum value. If the height difference is larger than a pre-defined height threshold, the pixel under examination is classified as a non-ground point. Otherwise, the slope difference between the current and the following pixel in the scanline direction will be calculated. If the slope is greater than a pre-defined slope threshold, this point is also labelled as a non-ground point. If the slope is less than the slope threshold and positive, the point is classified the same as the label of the previous point. If the slope is less than the slope threshold and negative, the point is labelled as a ground point. The process is repeated for all 4 scanlines in two directions leading to a total of 8. Finally, a pixel is classified as a ground point if the results of more than five labels indicated this point as a ground point. Otherwise, the point is classified as a non-ground point.

Beumier and Idrissa (2016) developed a DTM extraction method from DSMs based on photogrammetry. As a pre-processing step, the input DSM is smoothed using a Mean-shift filter. Then, the segmentation followed by region filtering are implemented and repeated. The segmentation technique is implemented for separating the DSM into regions based on the height information. The region filtering is applied for rejecting parts that are locally higher, which typically corresponds to non-ground objects such as buildings depending on neighbourhood analysis. The remaining regions are normally match roads, large surface or fields. Finally, holes resulted from rejecting non-ground objects in the previous step is then interpolated using bilinear interpolation technique to generate final DTM.

According to Meng et al. (2010) classification, the approach introduced by Beumier and Idrissa (2016) belongs to segmentation category. By contrast, our development approach is positioned in the context of directional scanning category. Hence, the next section will discuss the advantages and disadvantages of the directional scanning algorithms focusing on the MSD method because it belongs to the same class and in the context of DSMs based on photogrammetry.

3. METHODOLOGY

Based on the discussion of the main drawbacks of the MSD algorithm, we will introduce our NGPs method in detail, and will outline the different steps involve to improve its ability of overcome these drawbacks.

3.1 Problem Statement

Generally, three major steps are required to extract a DTM from a DSM following the MSD method (Perko et al., 2015):

1. Selecting pixels in the input DSM positioned on bare-ground areas.
2. Eliminating all other non-bare-ground areas.
3. Filling holes using an interpolation method.

Drawback 1: No ground points in the scanlines

According to (Perko et al., 2015), the most critical step of the approached is the determination of pixels that are positioned on the bare-ground areas. For each scanline within the extent of the filter area, a minimum of one point is required to be determined as belonging to the ground. The first drawback is given in the case that none of the points in the scanline is actually located on a bare-ground region is considered. However, this is easily possible and depends on the complexity of the structure in the area of interest, e.g. density of the built-up area, large building dimensions, very high resolution DSMs and the used filter size. If incorrect points are selected as minima, it will result in the incorrect classification of all other pixels within the filter size. Extending the filter size and therefore the length of the scanlines is not an ultimate solution to the problem because it can lead to some raised bare-ground pixels being incorrectly classified as a non-ground pixel, especially if the area of interest is not flat. Consequently, incorrectly choosing the minimal value will influence the DTM extraction negatively.

Drawback 2: Selection of correct minima in sloped areas

The second drawback is related to the local slope correction, or in other words, the determination of correct minima in sloped areas. (Perko et al., 2015) uses terrain slope fitting by applying a 2D Gaussian smoothing filter to solve this problem. However, such drawback cannot be overcome optimally in this way as the input DSM values were manipulated by the smoothing step. As a result, the incorrect choosing of minimal values could also occur in this case.

Drawback 3: Slope angle as a measure in high resolution DSMs

Assuming that the minimum point was determined correctly, the third drawback is given when the height difference w.r.t. the currently evaluated pixel is less than a pre-chosen vertical threshold in all directions. In this case, slope differences between the current and the following pixel in the scanline direction will be considered to make a classification decision whether the pixel is a ground or non-ground pixel. This drawback is especially present in high resolution DSMs as they are common when extracted from airborne images. Figure 1 shows the pixel under evaluation in the centre highlighted in grey with its neighbours in original DSM (Figure 1a also called *oDSM*) and in the smoothed DSM (in Figure 1b also called *sDSM*) after the 2D Gaussian filter was applied. The evaluated pixel in the centre should get the classification result “ground point” as can be seen by the height values in the figure. Let’s assume that the previous pixels are classified as ground pixels, the slope threshold is set to 30 degree, and the GSD equals 0.14 m. All of the assumptions are likely for a DSM of a city with some slopes. For the example in Figure 1a, the difference of the scanline pixel in the top left corner to the currently evaluated pixel in the centre (*oDSMDiff*) is:

$$oDSMDiff = 8.2 - 8.37 = 0.17m$$

And analogue for the smooth DSM *sDSMDiff* is:

$$sDSMDiff = 8.2933 - 8.2939 = -0.0006m$$

Calculating the difference of *oDSMDiff* and *sDSMDiff* gives us:

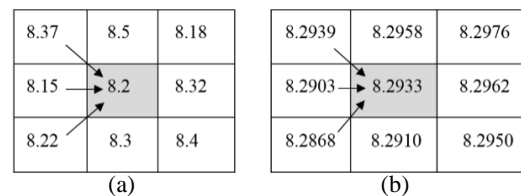
$$\Delta = oDSMDiff - sDSMDiff = -0.1694m \quad (1)$$

The calculation of the slope difference Δ based on Δ gives us:

$$\Delta = \text{atan}(\text{abs}(\Delta)/\text{GSD}) = 50.4251 \text{ degrees} \quad (2)$$

And hence, Δ is larger than the pre-defined threshold of 30 degrees indicating a non-ground point.

For the scanline running horizontally between the evaluated pixel and the left side, Δ is 0.047m and Δ is -18.5535 degrees. As Δ is smaller than the pre-defined threshold (and smaller than zero) the given label would be “ground point”. Similarly, the local slope Δ between the evaluated pixel and the bottom left scanline is 10.7155 degrees, and will be labelled similarly to the previous pixel “ground point” with Δ being smaller than the threshold. When continuing with calculating the slopes to all directions, the label “ground point” is given 3 times and the label “non-ground point” is given 5 times. Accordingly, the evaluated pixel in the centre will be classified as a non-ground pixel. Increasing the slope threshold to e.g. 60 degrees is not an optimal solution because 60 degrees corresponds to a height change equal to 24cm over a 14cm distance. These height changes are unlikely to appear everywhere without transitioning areas from ground to non-ground or vice versa. As a conclusion: The slope angle as a measure introduced in (Meng et al., 2009; Perko et al., 2015) seems to be not suitable for the DTM generation based on DSMs with a resolution of less than 1m GSD. For this reason the slope angle measure will be excluded from the processing of DTM extraction in our approach and will be replaced by an alternative solution.



Figures (1) represents the evaluated pixel under evaluation in the centre within the filter size with its 8 directions (a) showing the original DSM (*oDSM*) and (b) showing the smoothed DSM (*sDSM*).

3.2 Proposed DTM Extraction Approach

The drawbacks explained previously can be overcome by implementing the following enhancements:

1. Designing a new filter structure in such way that well-distributed ground points are selected.
2. Adding a vertical threshold for choosing minimum pixels to be accepted.
3. Building a network of accepted ground points as minima and storing them with their geo-reference positions of the original DSM in a 2D surface.
4. Creating an initial DTM using an interpolation method to fill the gaps between the created NGPs in the 2D surface.
5. Pixel to pixel comparison between the initial DTM and the original DSM for pairwise classification of ground and non-ground pixel by applying a second vertical threshold.
6. Finally, removing non-ground pixels and filling the remaining holes.

As one of the main contributions of the enhancement is the network of ground points, we will refer to the NGPs algorithm as our new proposed method.

First of all, if the input DSM contains outliers due to occlusion or mismatches, these outliers have to be removed. These areas and gaps then have to be filled with an interpolation method.

To overcome the first drawback, a new parameter is proposed to examine all points selected as minima in all executed scanlines. For example, for each executed scanline, there is one point selected as a ground point because it has a minimum value. The values of these points are sorted, and the first minimum value is eliminated in order to avoid points from too lower regions due to mismatches and the second value will be accepted instead. Afterwards, the height differences between the accepted point (second) and each of the remaining points are computed. If the height difference of a point is smaller than a pre-defined vertical threshold (1.1m in our case study), this point is also accepted.

Otherwise, it will be eliminated from the process because it might be a non-ground point.

To overcome the second and third drawbacks in a more reliable manner, a novel technique is proposed. In order to reduce the computation procedure we work on the input DSM (i.e. no smoothing is applied) to obtain well-distributed ground points (GP). Furthermore, in addition to the 4 scanlines proposed in the MSD approach (Figure 2a), another 4 scanlines are proposed given us a total of 8 scanlines (Figure 2b) which yielded 8 points. Pixels highlighted in an orange are excluded from the process to avoid scanlines sharing the same pixel as minimal values. The 8 points will be examined before they are accepted as a ground point as discussed previously. Figure 3a is shown of the original DSM which we estimated initial ground pixels from the previous step. Then, the accepted ground points are stored with their geo-reference position to build a network of ground points (NGPs) for one scanline. Further, an initial DTM is generated by filling the gaps between the created NGPs (Figure 3b). Afterwards, a vertical pixel to pixel comparison between the initial DTM and the original DSM is performed by applying a vertical threshold shown in figure 3c. If the height difference is less than the chosen vertical threshold (e.g. absolute 0.4m), this pixel is classified as a ground pixel, otherwise, it is a non-ground pixel. These classified ground pixels correspond to the ground mask shown as a red line in Figure 3d. Finally, the remaining gaps between the ground mask are filled through interpolation shown as a dash red in figure 3e. Because the pixel to pixel vertical height threshold between the initial DTM and the original DSM is a more reliable measure than the slope threshold, the NGPs algorithm is a powerful alternative solution.

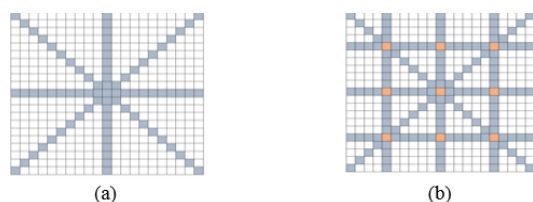


Figure (2). Directional scanlines for minimum point selection (a) MSD approach and (b) NGPs approach.

4. EVALUATION

4.1 Dataset

For the evaluation, the Vaihingen (Germany) dataset provided by the former ISPRS working group III / 4 is used. This test dataset was chosen because:

- the images were captured with high resolution airborne cameras,
- reference data in the form of airborne laser scanning (ALS) data are available; and
- the data represent a complex scene of a high dense urban area with many buildings, vegetation and cars as well as a slope of the terrain.

The specifications of the dataset are as followed:

- Airborne images (8 cm GSD) associated with their orientation parameters; acquired using the platform Intergraph/ZI DMC with 0.12 m focal length (Cramer, 2010). The colour information consists of three bands: near infrared (NIR), red (R), and Green (G). The derived true orthophoto mosaic is provided.
- Airborne Laser Scanning (ALS) data captured using a Leica ALS50 system with 4 points/m² density average and its derived DSM with 25 cm GSD.
- Digital surface models (DSMs) generated by dense matching using the Match-T software with 14 cm and 9 cm spatial resolution.

From this dataset, we especially focus on area 1 (“Inner City”) and area 2 (“High Riser”). While area 1 is especially suitable

due to the sloped terrain covered with a complex irregular building structure including vegetation, area 2 was selected because of the existence of high raised buildings with larger objects located on top of some of those buildings.

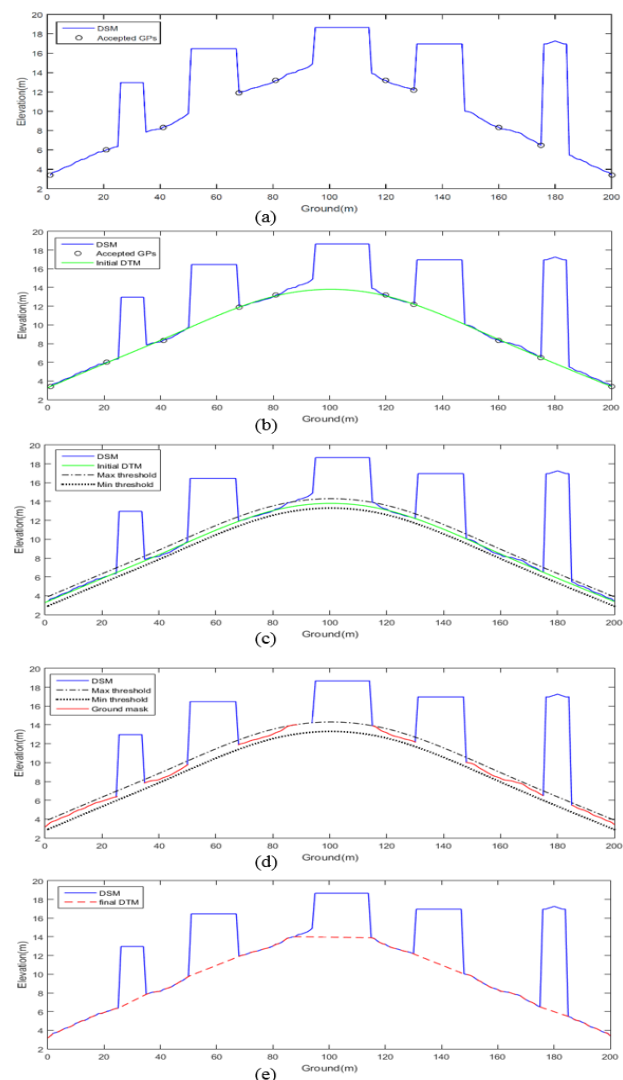


Figure (3). Logical steps of DTM generation by NGPs algorithm. Ground points (GPs) selected as minima on an artificial DSM (a), initial DTM (b), and vertical threshold limits (c), ground mask (d), and final DTM (e).

4.2 Parameter settings

The parameters used for the MSD approach are provided in Table 1, and for the NGPs approach accordingly in Table 2; the same parameters were used for both areas. The values for the filtering window size are fixed to 53m for both approaches and both tests. The filtering window size is based on the dimensions of the buildings in the scene. For the MSD approach, the height threshold is based on the absolute height difference between ground and buildings while the slope threshold is based on the height difference between the centre pixel and the one in scanline direction as well as on GSD. Regarding the NGPs approach, the first vertical threshold is based on the height difference between ground points selected as a minimum from different scanline. For non-flat areas, the value for this threshold should be increased by increasing the filtering window size and the angle of slope and vice versa. The second vertical threshold is based on the vertical height difference between the initial DTM and the original DSM. Increasing this threshold means capturing higher regions located between ground and non-ground regions.

Filter size	53 m
Height threshold	3 m
Slope threshold	30

Table 1. Parameters' values (MSD approach).

Filter size	53 m
Vertical threshold for accepting ground points	1.1 m
Vertical threshold for detecting ground mask	0.4 m

Table 2. Parameters' values (NGPs approach).

For the creation of the nDSMs for both approaches and test sets a threshold of 2m is applied, i.e. only objects with a height of more than 2m is shown in the nDSMs.

4.3 Qualitative Evaluation

4.3.1 Area1: Figure 4 shows DSM derived from ALS data (a) and a reference DTM (b). The resulting DTMs using the MSD algorithm and the NGPs algorithm are presented in Figure 5. The Figure shows the ortho image for Area 1 (a), the DSM with 14cm GSD based on image matching (b), the detected ground regions mask (c), the resulting DTM (d) and the created nDSM (e) using the MSD algorithm as well as the network of ground points accepted as minima values (f), the initial DTM (g), the detected ground regions mask (h), the extracted DTM (i) and the resulting nDSM using the NGPs algorithm (j).

Obviously, when comparing the results of the ground masks (Figures 5c and h), details are lost by MSD approach as highlighted with yellow circles (Figure 5c), while larger regions belonging to bare-ground are successfully detected by the NGPs approach (Figure 5h). The NGPs approach is able to segment buildings, high vegetation, and even some cars, and excluded them from the ground mask.

Furthermore, the second last row of images in Figure 5 show the resulting DTMs. There is one highlighted area in the generated DTMs of the MSD approach (Figures 5d) as well as in the NGPs approach (Figure 5i). This area indicates an error in the generated DTM whereas instead, it is actually due to an error in the input DSM. The cause of this error is unknown but can be seen more clearly in Figure 4. While in the LiDAR dataset a building within the highlighted area is clearly visible (Figure 4a), this area is classified as terrain in the DTM processed by LASTool (Figure 4b).

Lastly, the normalised DSMs (also called nDSMs) are presented in the last row of Figure 5. The nDSMs are created by subtracting DTM from DSM. Then, by thresholding the nDSM (above 2m), all non-ground point, i.e. buildings and high vegetation will be obtained. The differences are clearly visible and are highlighted in the figure. While many raised bare-ground regions were not detected by the MSD approach leading to a cluttered nDSM (Figure 5e), the nDSM created by NGPs represents the location of buildings more realistic (Figure 5j). This is mainly due to the successful identification of ground points (Figure 5f).

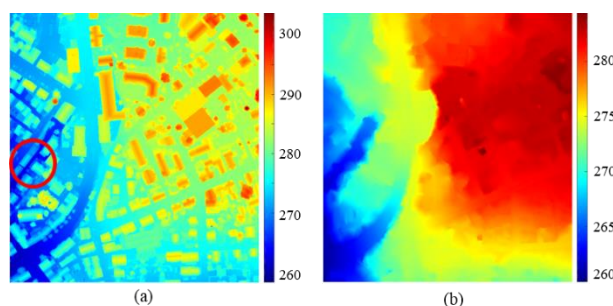


Figure 4. Area1 DSM derived from LiDAR data (a) and the reference DTM (b).

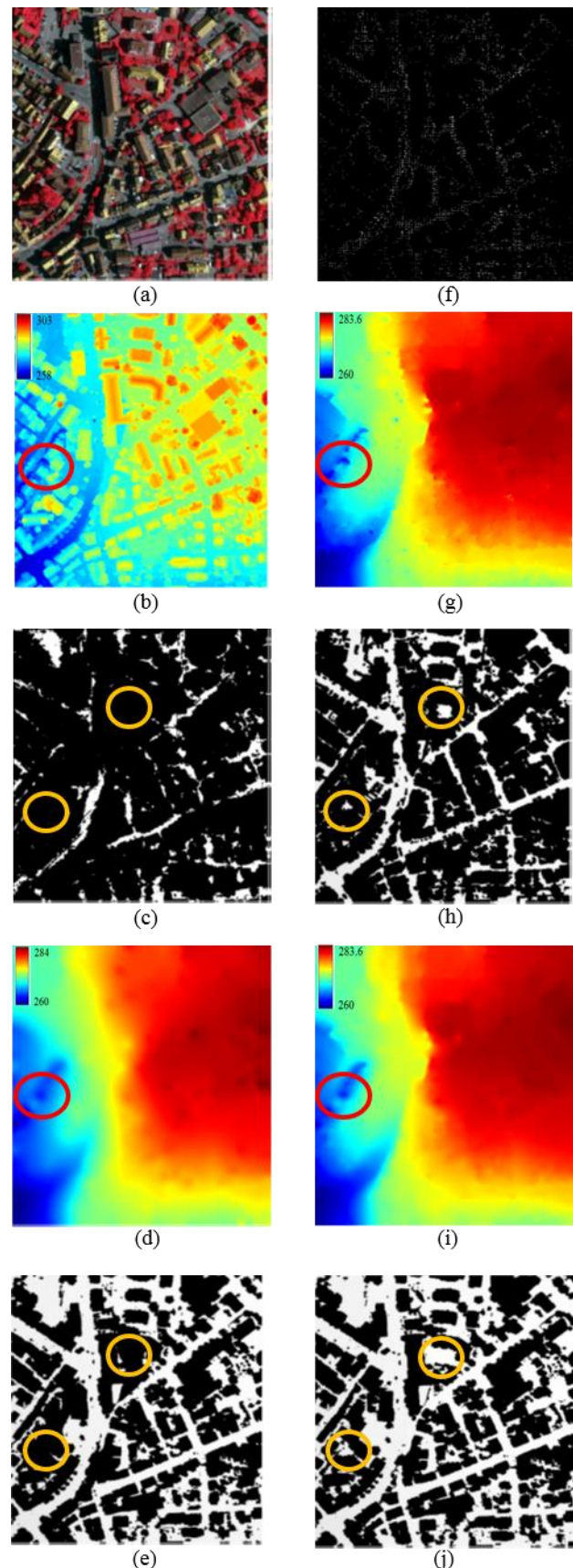


Figure 5. Ortho image for area1 (a) and input DSM (b). Ground mask (c), DTM (d), and nDSM (e) by MSD method. Network of ground points (f), initial DTM (g), ground mask (h), final DTM (i), and nDSM (j) by NGPs method.

4.3.2 Area2: DSM derived from ALS data (a) and a reference DTM (b) are shown in Figure 6. Figure 7 presents the outputs of both algorithms. The Figure shows the ortho image for Area 2 (a), the DSM of 9cm GSD based on image matching (b). The resulting ground masks, DTMs and nDSMs are in the same order as for area 1.

The difference in the created ground masks of the MSD approach (Figure 7c) to the NGPs approach (Figure 7h) is clearly visible. For instance, wide ground regions have been lost from the ground mask using the MSD approach and non-ground points are clustered together. In contrast, our proposed algorithm NGPs successfully detects those areas.

Figures 7d and 7i show the DTMs created by MSD and NGPs approaches respectively. Both DTMs seem to be similar except the areas highlighted with red circles. NGPs DTM values in this area are clearly higher than the DTM created by MSD and even higher than LiDAR DTM as highlighted in figure 6b. In fact, the true height value for this ground area is higher than all created DTMs, as visible in the ortho image (Figure 7a) and in the DSM (Figure 7b). That means, DTM created by NGPs is the closest to the true value in this highlighted area.

In spite of the significant improvements in the quality of the ground mask created by NGPs approach, the created DTMs (Figure 7d and 7i) and nDSMs (Figure 7e and 7j) from MSD and NGPs approaches look similar. This is due to the topographic surface of area 2 being nearly flat. The one difference which can be seen is highlighted with red circles. The high difference is up to 6 meters with a sudden change. Such case is very difficult to classify correctly because usually large height changes are used to actually distinguish between ground and non-ground regions. However, a smaller part in this area is incorrectly classified as non-ground region by NGPs (Figure 7j) in comparison to MSD (Figure 7e).

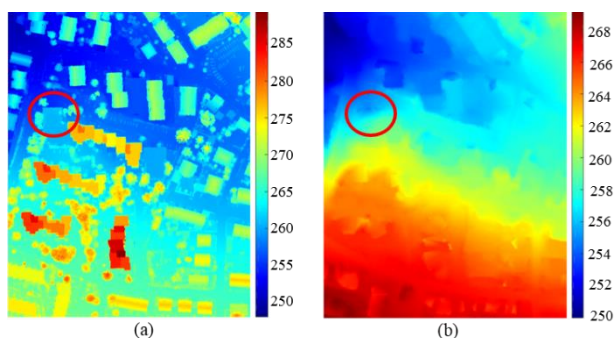


Figure 6. Area1 DSM derived from LiDAR data (a) and the reference DTM.

4.4 Quantitative Evaluation

4.3.1 Area 1

For the purpose of the quantitative evaluation, a reference DTM is created using LAsTools software and subtracted from the DTMs created by the MSD and the NGPs methods (Figure 8). For both difference images, positive differences means created DTMs higher than LiDAR DTM and vice versa. For some areas the height difference reaches up to 4 m (red areas) which is too high. However, please note that the error inside of the areas highlighted with circles is related to the error in the original DSM as discussed earlier. The second error which is marked by red arrows is related to the interpolation technique used in both methods. While inward interpolation has been used in the MSD and NGPs method, the LAsTools uses standard linear interpolation.

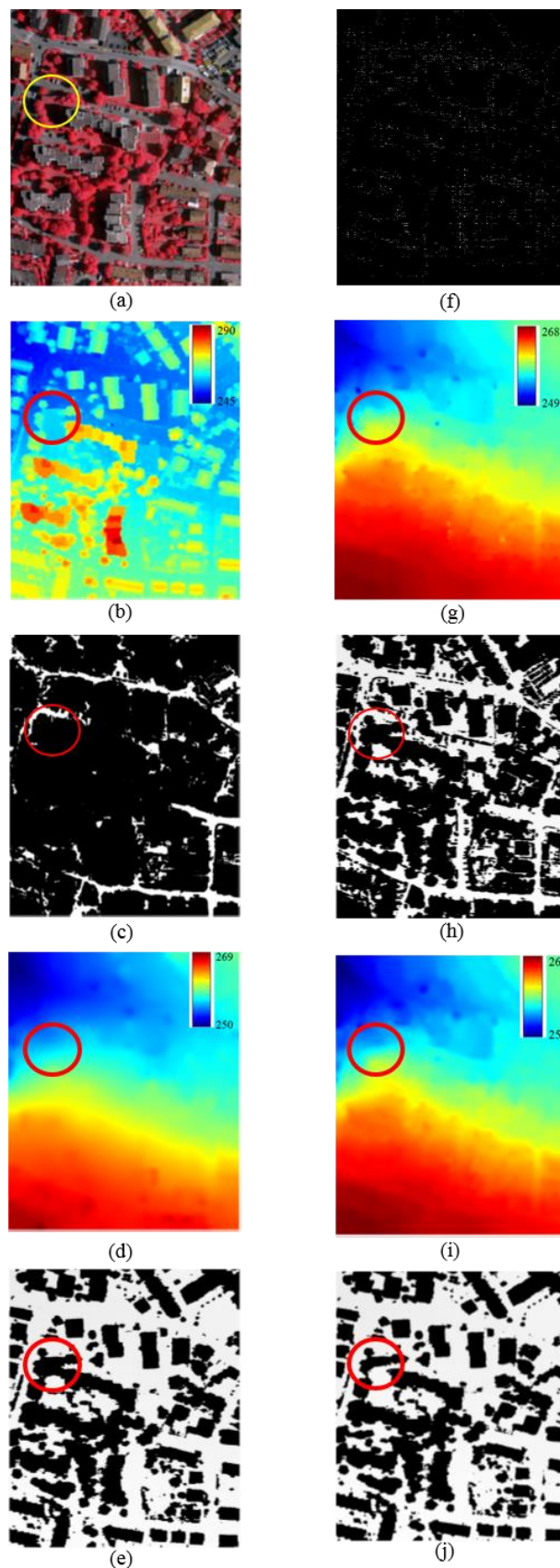


Figure 7. Ortho image for area1 (a) and input DSM (b). Ground mask (c), DTM (d), and nDSM (e) by MSD method. Network of ground points (f), initial DTM (g), ground mask (h), final DTM (i), and nDSM (j) by NGPs method.

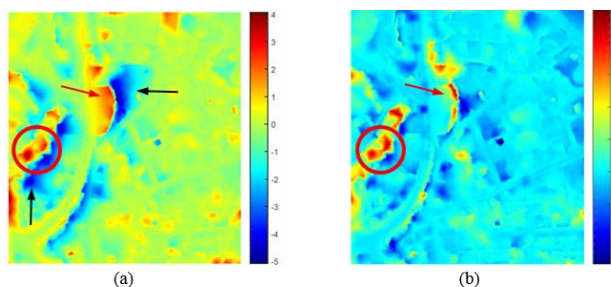


Figure 8. Area (1) height difference maps of the LiDAR DTM compared to the MSD (a) and the NGPs (b).

The negative height differences are highlighted by black arrows in Figure 8 and are only presented in the DTM created by the MSD approach. Those differences are up to 4m and compared with the 2m by the NGPs approach will have to be flagged as gross errors. The major reason is that raised ground regions are lost from the ground mask as discussed in Figure 5c. Hence, the created DTM is interpolated under the ground instead of on the ground and therefore are highlighted as incorrect classified. Based on the difference DTMs and after excluding the gross errors, the mean-square error (MSE) and the standard division (SD) are computed and presented in Table 3.

	MSE	SD	Computation time
MSD	0.8560	0.3660	556.06 s
NGPs 4d	0.4574	0.2666	50.19 s
NGPs 8d	0.3814	0.2492	52.60 s

Table 3. Statistics of mean-square error (MSE) and standard division (SD) of the height differences between MSD and NGPs compared with the LiDAR DTM as well as the time required to execute the algorithms.

For this evaluation step, we used the NGPs algorithm with different number of scanline directions: 4 (NGPs 4d) and 8 (NGPs 8d). While the NGPs algorithm should be run with 8 directions, 4 were also used in this experiment in order to evaluate how much the NGPs improves the results compared to the MSD approach by only using a different approach of determining the ground initially. Hence, we can analyse the impact of the successfully detected ground points and their distribution. The mean squared error decreases from the MSD to the NGPs 4d and then further to the NGPs 8d. Therefore, we can conclude to that the selection of the ground points improves the results. However, the introduction of additional scanlines seems to have a higher impact as the drop of the mean squared error is higher. This conclusion is also verified when looking at the standard division. Furthermore, the computation time required to execute the NGPs algorithm is significantly less than the MSD algorithm due to the reduced complexity as discussed previously.

4.3.2 Area 2

The height difference maps of the created reference DTMs compared to the MSD method (Figure 9a) and the NGPs method (Figure 9b) are both significantly better than the DTMs created for Area 1 (Figure 8). This is mainly due to the fact that the topographic surface is mostly flat, and due to that there are no visible errors in the original DSM. The highlighted area (red circle) in Figure 9a indicates large negative errors from up to 2m in the MSD extracted DTM. The MSD approach is still facing the same challenge as highlighted and discussed previously in Figure 8, and hence confirming previous outcomes. In contrast, the maximum negative error in this area in the DTM created by the NGPs method is smaller by approximately 0.5m. While there are no significant lower sections in the NGPs there is one higher area by nearly 2.5m as shown in Figure 9b. This area is also highlighted previously in Figures 6a and 6b. In fact, the correct value for this area is higher than what our NGPs approach determines. Consequently,

NGPs is significantly better and therefore the DTM values are the closest to the truth values.

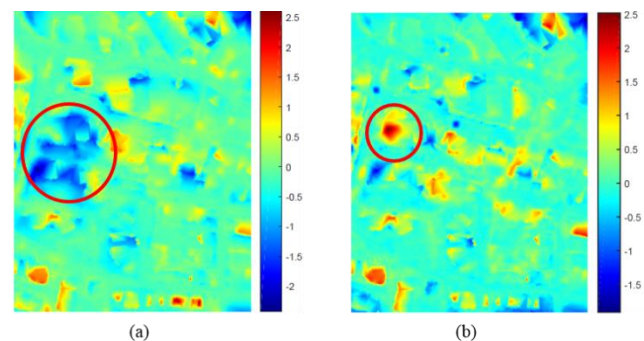


Figure 9. Area (2) height differences maps of the created DTMs with LiDAR DTM. MSD (a) and NGPs (b).

Table 4 shows the calculated Mean-squared errors (MSEs) and standard division (SD) of the height differences. The MSE and SD equal to 0.1942 and 0.1079 for DTMs created by MSD, 0.1488 and 0.0873 for (NGPs 4d), and 0.1348 and 0.0879 for (NGPs 8d) respectively. Accordingly, the quality of the created DTM by NGPs is slightly improved in area 2 compared to area 1 due to the fact that area 2 is rather flat. However, similar to area 1, NGPs algorithm requires significantly less time as concluded previously.

	MSE	SD	Computation time
MSD	0.1942	0.1079	5480.5 s
NGPs 4d	0.1488	0.0873	319.47 s
NGPs 8d	0.1348	0.0879	327.16 s

Table 4. Statistics of mean-square error (MSE) and standard division (SD) of height differences between MSD and NGPs compared with the LiDAR DTM as well as the time required to execute the algorithms.

5. CONCLUSION

This paper presents a simple and powerful filtering algorithm for the DTM extraction from airborne stereo images derived DSMs. This algorithm is an enhancement of the MSD approach proposed by (Perko et al., 2015). In contrast, to the original approach, the newly proposed NGPs approach solves the local slope problem in a more reliable way and with less complexity using reliable and well distributed initial ground points. A further extension is the increase of the number of scanlines. Hence the quality of the generated DTMs and further the nDSMs are significantly improved.

Two different datasets have been used to evaluate the NGPs method, and to compare the performance to the MSD method. In these testing similar filter size for both algorithms were used. The resulting DTMs were evaluated using qualitative and quantitative measurements. The visual inspection, as well as the objective measurements of the mean square error and the standard deviation, confirmed the efficiency and the robustness of the NGPs approach compared to the MSD approach.

However, while the initial DTMs created by our NGPs approach are quite acceptable for certain applications, the introduction of further processing technique maybe required. The goal of those techniques are to simplify and therefore to speed up the processes of the NGPs selection. For instance, it is not necessary to move the filter one by one pixel along the x and y directions over the whole DSM. First experiments show that moving 5 pixels in both directions will very likely yielded similar results but will require less computing time. Furthermore, eliminating very small regions from the created ground regions mask could be a useful option for enhancing the

created DTM accuracy because they might be incorrect (Perko et al., 2015).

In addition, while the inward interpolation technique is so far used for filling holes in the ground mask produces satisfying results, the finally created DTM could be also smoothed by using an average, median, or any other smoothing filter to obtain smoother DTM surfaces.

ACKNOWLEDGEMENTS

This work was supported by The Higher Committee for Education Development (HCED) in Iraq. Thanks to Dr Roland Perko for his cooperation and providing his matlab source code. The Vaihingen data set was provided by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) [Cramer, 2010]: <http://www.ifp.uni-stuttgart.de/dgpf/DKEP-Allg.html>.”

REFERENCES

- Arefi, H., d'Angelo, P., Mayer, H., & Reinartz, P. 2009. Automatic generation of digital terrain models from cartosat-1 stereo images. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXVIII-1-4-7/W5, pp. 1-6.
- Beumier, C., & Idrissa, M. 2016. Digital terrain models derived from digital surface model uniform regions in urban areas. *International Journal of Remote Sensing*, 37(15), pp. 3477-3493.
- Bulatov, D., Häufel, G., Meidow, J., Pohl, M., Solbrig, P., & Wernerus, P. 2014. Context-based automatic reconstruction and texturing of 3D urban terrain for quick-response tasks. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93, pp. 157-170.
- Chaabouni-Chouayakh, H., Arnau, I. R., & Reinartz, P. 2013. Towards automatic 3-D change detection through multi-spectral and digital elevation model information fusion. *International Journal of Image and Data Fusion*, 4(1), pp. 89-101.
- Cramer, M. 2010. The DGPF-test on digital airborne camera evaluation—overview and test design. *Photogrammetrie-Fernerkundung-Geoinformation*, 2010(2), pp. 73-82.
- Förstner, W. 1982. On the geometric precision of digital correlation. *Int. Arch. Photogrammetry & Remote Sensing*, 24(3), pp. 176-189.
- Haralick, R. M., Sternberg, S. R., & Zhuang, X. 1987. Image analysis using mathematical morphology. *IEEE transactions on pattern analysis and machine intelligence*, (4), pp. 532-550.
- Hirschmuller, H. 2008. Stereo processing by semiglobal matching and mutual information. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30, pp. 328-341.
- Krauß, T., Arefi, H., & Reinartz, P. 2011. Evaluation of selected methods for extracting digital terrain models from satellite born digital surface models in urban areas. *International Conference on Sensors and Models in Photogrammetry and Remote Sensing (SMPR 2011)*, pp. 1-7.
- Krauß, T., & d'Angelo, P. 2011. Morphological filling of digital elevation models. *Proc. ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* XXXVIII-4/W19, pp. 165-172.
- Krauß, T., d'Angelo, P., Kuschik, G., Tian, J., & Partovi, T. 2015. 3D-Information fusion from very high resolution satellite sensors. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-7/W3, pp. 651-656.
- Krauß, T., Lehner, M., & Reinartz, P. 2008. Generation of coarse 3D models of urban areas from high resolution stereo satellite images. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 37, pp. 1091-1098.
- Krauß, T., & Reinartz, P. 2010. Urban object detection using a fusion approach of dense urban digital surface models and VHR optical satellite stereo data. *Int. Arch. Photogrammetry & Remote Sensing*, 39, pp. 1-6.
- Meng, X., Currit, N., & Zhao, K. 2010. Ground filtering algorithms for airborne LiDAR data: A review of critical issues. *Remote Sensing*, 2(3), pp. 833-860.
- Meng, X., Wang, L., Silván-Cárdenas, J. L., & Currit, N. 2009. A multi-directional ground filtering algorithm for airborne LIDAR. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(1), pp.117-124.
- Perko, R., Raggam, H., Gutjahr, K., & Schardt, M. 2015. Advanced DTM generation from very high resolution satellite stereo images. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1, pp.165-172.
- Tian, J., Krauss, T., & Reinartz, P. 2014. DTM generation in forest regions from satellite stereo imagery. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-1, pp. 401-405.